BIOLOGICALLY INSPIRED COMPUTATIONAL NEURAL MODELS FOR MOTIVATED BEHAVIOR, LEARNING, AND MEMORY

BY

EKATERINA DMITRIEVNA GRIBKOVA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Neuroscience
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Professor Rhanor Gillette, Chair and Director of Research
Professor Martha U. Gillette
Associate Professor Daniel A. Llano
Associate Professor Prashant G. Mehta

ABSTRACT

The fields of artificial intelligence (AI) and machine learning have vastly expanded in the past decade, with a variety of modern applications, ranging from computer vision to language processing and medical diagnostics. While the majority of AI applications involve data classification, detection, and predictive modeling, fewer studies have explored the creation of motivated autonomous agents. The integration of neurobiological principles into AI, such as mechanisms involved in dopaminergic reward learning circuits, has been crucial for advancing more natural and biologically plausible forms of AI. The goal of this thesis is to introduce a set of biologically inspired models for motivated behavior, learning, and memory, that can be incorporated into artificially intelligent agents and networks. These models may also provide insights into the biological processes of episodic memory, aesthetics, and complex cognitive processes, as well as their evolution.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Professor Rhanor Gillette, for making so much of my work possible and for encouraging me to follow my dreams. He is the greatest advisor, supporter, and friend I could have ever asked for. I am deeply grateful to my previous advisor, Professor Dan Llano, for giving me the chance to pursue my first computational project starting as an undergraduate student, mentoring me in neurophysiology, and further supporting me throughout my graduate career. I would like to thank Professors Martha Gillette, Prashant Mehta, for being part of my thesis committee, Professor Mark Nelson as a former committee member, as well as Doctor Sam Beshers and Stephanie Pregent, all of whom have provided me with incredible support and invaluable advice over the years.

The 2017 Methods in Computational Neuroscience course at the Marine Biological Laboratory in Woods Hole was an unforgettable experience that led me to pursue my computational projects further, so I thank the 2017 course directors, Professors Mark Goldman and Michale Fee, the many professors, and my fellow students, who made the course so special.

Lastly, I am truly grateful to all of my friends, including those in the Neuroscience program, from the Gillette and Llano labs, from my undergraduate studies, the MCB Honors program, and especially my family, for always being there for me.

Dedicated to my parents

TABLE OF CONTENTS

CHAPTER 1: OVERVIEW

The fields of artificial intelligence (AI) and machine learning have vastly expanded in the past decade, with a variety of modern applications, ranging from computer vision to language processing and medical diagnostics. While the majority of AI applications involve data classification, detection, and predictive modeling, fewer studies have explored the creation of motivated autonomous agents. The integration of neurobiological principles into AI, such as mechanisms involved in dopaminergic reward learning circuits, has been crucial for advancing more natural and biologically plausible forms of AI. The goal of this thesis is to introduce a set of biologically inspired models for motivated behavior, learning, and memory, that can be incorporated into artificially intelligent agents and networks. These models may also provide insights into the biological processes of episodic memory, aesthetics, and complex cognitive processes, as well as their evolution.

Chapter 2 introduces the first model, Algorithm of Selectivity by Incentive, Motivation and Optimized Valuation (ASIMOV), an agent-based simulation that shows how a simple aesthetic sense and addiction dynamics can emerge from the integration of homeostatic reward circuit into a neuronal network for cost-benefit foraging decisions. In **Chapter 3**, Learning and Memory, ASIMOV is elaborated with the addition of the Feature Association Matrix (FAM), which enables the formation of simple episodic memory, allowing the ASIMOV agent to create spatial maps of its environment that it can use to maximize the rewards that it obtains in foraging. The FAM is an abstraction of physiological circuits implicated in episodic memory, such as the auto- and hetero-associative circuits of the hippocampus, and one of its learning rules is based on the timing between two stimuli. While learning in hippocampal circuits clearly depends on the temporal sequence of presented stimuli, it is unknown if single neurons are able

to learn and alter the timing of their inputs. Therefore, in **Chapter 3**, I introduce a model related to the FAM, the Synaptic Input Time Difference Learning (SITDL) model, which may explain how individual synapses use NMDA receptors to learn the timing difference between glutamate input and dendritic voltage input. This would be especially useful for maturation of synapses that rely on extreme temporal precision, as well as for memory formation. While the existence of such a mechanism is yet to be explored in physiological networks, SITDL may be used in artificial neural networks (ANNs) with more complex and realistic neuron models, in contrast to learning rules found in the FAM. **Chapter 4** contains supplementary materials on the mutual information estimator, Adaptive partition using Interspike intervals MI Estimator (AIMIE), used in Chapter 3 for analysis. **Chapter 5** provides a brief conclusion on models introduced in Chapters 2-3, which utilize key neurobiological and behavioral concepts, and may provide insight into evolution and development of more complex behavior and memory, as well as significant enhancement of AI agents and algorithms.

CHAPTER 2: MOTIVATED BEHAVIOR

Behavior-oriented AI is a scientific discipline that examines the emergence of behavior, and how it becomes intelligent and adaptive (Steels, 1993). In this field, the problem of intelligence is presented within the general context of biology, with intelligent behavior specifically defined as behavior that maximizes an agent's survival in its environment. As such, there is a focus on the principles which can be formulated at the behavioral level, rather than the actual physical basis of behavior. Behavior-oriented AI researchers often employ physical and virtual agents, typically using ANNs, subsumption architectures, combinatorial circuits, and dynamical systems to design and implement agents' behavior programs (Steels, 1993). Of particular interest are AI agents that show motivated, goal-directed behavior. One of the earliest examples is Tolman's schematic Sowbug, which is considered to be the first prototype that described behavior-based robotics architecture, prior to subsumption architectures (Arkin, 2003; Tolman, 1939). While there has been significant work in the creation of motivated agents since Tolman's studies, it seems that few have explored this in the context of evolution of more complex behaviors.

ASIMOV[1]

A rudimentary aesthetic sense is found in the stimulus valuations and cost-benefit decisions made by primitive generalist foragers. These are based on factors governing personal economic decisions: incentive, appetite, and learning. We find that the addictive process is an

[1] *Scientific Reports Permissions Policies: https://www.nature.com/nature-research/reprints-and-permissions/permissions-requests*. **Chapter contains previously published material from:** Gribkova, E. D., Catanho, M., & Gillette, R. (2020). Simple Aesthetic Sense and Addiction Emerge in Neural Relations of Cost-Benefit Decision in Foraging. *Scientific Reports, 10*(1), 1-11.

extreme expression of aesthetic dynamics. An interactive, agent-based model, ASIMOV, reproduces a simple aesthetic sense from known neural relations of cost-benefit decision in foraging. In the presence of very high reward, an addiction-like process emerges. A drug-like prey provides extreme reward with no nutritive value, initiating high selectivity and prolonged cravings for drug through reward learning. Varying reward experience, caused by homeostatic changes in the neural circuitry of reward, further establishes the course of addiction, consisting of desensitization, withdrawal, resensitization, and associated changes in nutritional choice and pain sensitivity. These observations are consistent with the early evolution of addiction mechanisms in simple generalist foragers as an aesthetic sense for evaluating prey. ASIMOV is accessible to inspection, modification, and experiment, is adaptable as an educational tool, and provides insight on the possible coevolutionary origins of aesthetics and the addiction process.

INTRODUCTION:

The aesthetic sense is a subjective, evaluative faculty used to distinguish positive and negative qualities of situations, objects, and constructs, and to bias behavioral decision toward or away from those stimuli. It is based on built-in preferences and feature detection, as well as learned preferences established from experience through reward learning. In humans, the highly developed aesthetic sense extends from judgements of taste and beauty to disgust. In other animals, it notably functions in mate choice, nest building, and foraging.

Darwin and others (Darwin, 1888; Prum, 2017) attributed the origin of the aesthetic sense to mate choice and reproductive displays, as are notable in many vertebrates. However, here we explore the ramifications of a potentially earlier origin in the foraging decisions of generalist animal species, where valuations of potential prey are made in estimates of nutritional value that

factor in need, learned attributes, and risk. A primitive basis of the aesthetic sense appeared in our studies of the neuronal circuitry of decision in the predatory, generalist sea slug, *Pleurobranchaea californica*, in the animal's ability to evaluate stimuli in contexts of motivation and reward learning. It was initially implemented in an agent-based foraging simulation, Cyberslug (Brown et al., 2018). That agent made foraging decisions for approach or avoidance like the real animal, based on stimulus quality, motivation, and reward learning, and satisfied requirements of optimized foraging.

We introduced the original simulation as an example of simple neuronal relations that could be elaborated for more complex cognition and behavior, as may have happened to ancestral bilaterians in evolution. Accordingly, here we introduce a new version, ASIMOV, which is upgraded for more realistic expression of aesthetic sense with 1) explicit representations of dynamic hedonic tone in reward experience and of noxious pain with direct accesses to the agent's appetitive state, and 2) a novel mechanism of homeostatic plasticity that contributes to use-dependent desensitization of the reward experience. ASIMOV implements two key forms of plasticity characteristic of the natural aesthetic sense. First, reward learning, which can establish complex preferences to guide acquisitive, synthetic, and creative behaviors. Second is use-dependent habituation to the reward experience produced by repeated exposure to a stimulus. For instance, for animals given access to an unlimited supply of a new and highly palatable food, the relative palatability of that food may decline (Young, 1946). Congruently, attention to acoustic and visual stimuli is balanced between repetition/regularity and novelty (Berlyne, 1971).
We found in the simulation that the dynamic process of addiction to high reward stimuli emerges as an extreme expression of aesthetic dynamics. Addiction begins with simple reward learning, but the course of addiction through desensitization, withdrawal, and resensitization, with

associated changes in nutrition and pain sensitivity, is established through homeostatic changes in the neural circuitry that expresses reward experience. ASIMOV, with its simple homeostatic reward circuit, also reproduces the dynamics of earlier important models of addiction, such as the opponent-process model (Solomon & Corbit, 1974) where the direct reward input stimulation in ASIMOV can be considered as a primary hedonic process, and the homeostatic plasticity in its reward circuit is analogous to the opponent hedonic process. The results support the view that addiction involves unusually large, rewarding stimuli, to which the forager is not adapted and becomes impaired in its volition.

The ASIMOV model is broadly accessible to inspection, modification, and experiment, is easily adaptable as an educational tool, provides insights on the possible evolutionary origin of the addiction process, and may be developed further for more complex aesthetic tasks.

METHODS:

ASIMOV (Algorithm of Selectivity by Incentive, Motivation and Optimized Valuation) derives from a previous simulation (Brown et al., 2018) based on reward learning and motivation, and founded on neuronal relations used in cost-benefit choices of foraging by the predatory sea-slug *Pleurobranchaea*. A simple aesthetic sense is expanded in ASIMOV with a homeostatic reward circuit expressing reward experience. Simulating the aesthetic sense requires mechanisms for preference formation by reward learning, and for initiation and terminating preference-seeking behaviors by modulating appetitive state via incentive, reward experience, and satiation. With these, the addiction process emerges in an extreme expression of aesthetic preference.

The forager encounters two virtual prey in the environment, the benign Hermi and noxious Flab (based on natural prey of *Pleurobranchaea*), and a high-reward Drug that provides no nutrition. Each prey and Drug secrete their own signature odors. The ASIMOV forager's simple aesthetic sense is altered with reward learning, as by experience it associates different signature odors with positive and negative expected rewards, establishing a set of dynamic preferences. The addiction process occurs as exaggerated preference for a high-reward item with a specific odor. For the forager, odor signature is the context in which drug is acquired, analogous to place preference in humans for the site where a drug is obtained, how it is ingested, or the company of like-minded acquaintances. ASIMOV is a minimalist model without critical conditional statements, in which decision emerges at thresholds attained by interactions of variables. This approach is aimed to conservatively model the primitive functionality of aesthetic sense in a simple forager.

ASIMOV Model Architecture

ASIMOV's structure is shown in **Figure 2.1**. The core of the model is the origin of behavioral choice in appetitive state, which controls economic decisions of foraging (Brown et al., 2018). Appetitive state represents the animal's biases towards appetitive behaviors, including prey tracking, handling, and consumption. By integrating reward experience, incentive, pain, and motivation, appetitive state controls the choice of an approach or avoidance turn, such that low appetitive state causes aversive responses to stimuli, and increasing appetitive state inverts turn response direction to one of approach (Gillette et al., 2000). Thus, appetitive state sets sensory thresholds for approach turns toward prey and subsequent feeding responses; in switching,

excitatory sensory input is routed from one side of the turn network to the other, to cause a turn towards the stimulus.

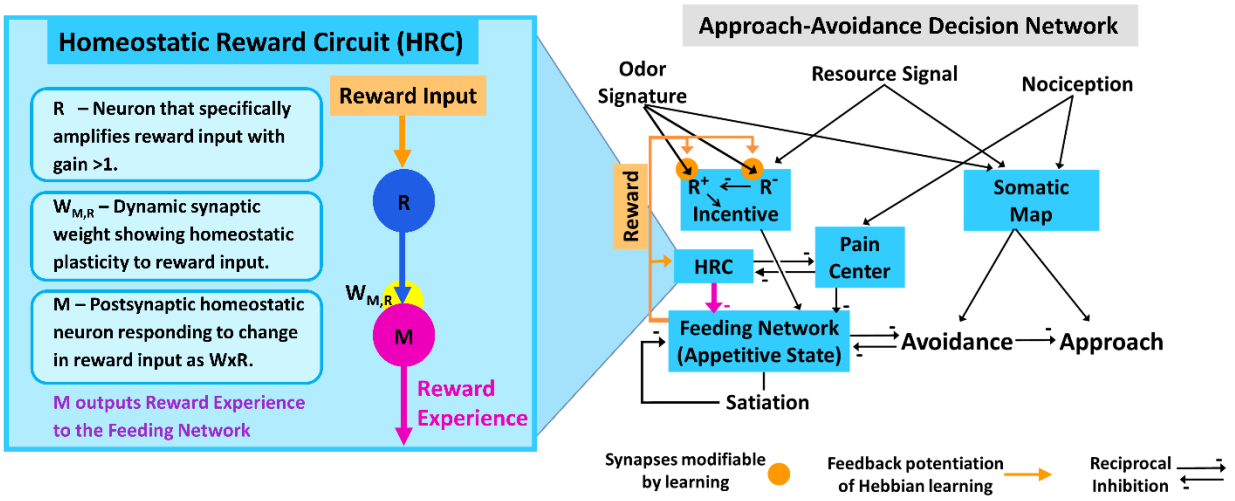

Figure 2.1. ASIMOV's neural network of foraging decision. Right: In the modified decision network from Cyberslug (Brown et al., 2018), Appetitive State (feeding network excitation) sums intrinsic and learned stimulus values as Incentive with motivation (Satiation) to regulate tur regulate turn response direction. In parallel, a map of stimuli (Somatic Map) from the animal's oral veil sets turn amplitude. Incentive sums sensory inputs predicting nutritional value (Resource Signal) with learned positive and negative values of prey odor signatures (R+ and R-), and is then integrated with motivation, Reward Experience, and Pain into Appetitive State (Feeding Network excitation). Positive and negative classical learning occur by outputs from the feeding network operating in approach or avoidance modes, respectively. Left: The Homeostatic Reward Circuit (HRC) mediates habituation to rewarding cues, the basis of Drug desensitization and withdrawal. HRC integrates rewards from prey consumption as Reward Experience and reduces Appetitive State in negative feedback. HRC comprises two connected rate-based neurons, R and M. R receives and amplifies reward output from Feeding Network. Neuron M expresses homeostatic plasticity, habituating to reward. M's activity is a product of the dynamic synaptic weight W and neuron R's activity. HRC activity lies in reciprocal inhibition with the Pain Center output; higher levels of pain suppress Appetitive State and cause aversion to painful stimuli. Thus, pain's suppressive effect is reduced by positive reward output from HRC.

Sensory inputs are: 1) a resource odor signal predicting nutritional content; 2) specific odor signatures for different prey species; 3) a place code for the averaged site of sensory input to the sensors (Somatic Map); and 4) nociception. (1) and (2) are summed as Incentive for resource and learned positive and negative values of prey odors (R^+ and R^-, respectively).

Incentive is then integrated with motivation, reward experience, and pain as Appetitive State in the Feeding Network. The sensory place code in Somatic Map acts as a template for turn response amplitude. Positive or negative learning are consequences of feedback from the feeding network operating in feeding or avoidance modes, respectively. Outputs of Appetitive State are reward and a converting function that switches the turn response to stimuli from avoidance to approach.

Reward experience is the output of a homeostatic reward circuit (HRC) module (**Figure 2.1**, left), resulting from reward circuitry activation, that feeds back to Appetitive State. HRC actions resemble habituation (McSweeney & Murphy, 2009). Reward input, as from a recreational drug, is amplified by neuron R and fed to postsynaptic neuron M, whose activity is the product of synaptic weight W and neuron R activity. The synaptic weight W between neurons R and M changes dynamically based on both presynaptic and postsynaptic activity, as well as baseline activity. With repeated or long enduring large reward stimuli, as in Drug reward, homeostatic plasticity desensitizes neuron M's response, which reduces positive reward effects (such as Drug reward) and causes them to decay faster. Thus, reward experience, a function of neuron M activity, differs with consumption of different prey, and changes drastically with rewarding Drug or withdrawal. If intake of rewarding prey or Drug is relatively frequent, reward experience diminishes due to homeostatic plasticity. Negative reward experience results from noxious prey consumption, and more severely, from Drug withdrawal.

Pain suppresses appetitive state, biasing decision towards avoidance. Pain and reward experience are opponent processes (Elman & Borsook, 2016) that are reciprocally inhibitory (**Figure 2.1**, left). Thus, reward experience also influences appetitive state by gating pain input. Each by itself at high values suppresses appetitive state; when either mode dominates, it becomes the major

suppressor of appetitive state. Positive reward experience attenuates aversive responses by opposing suppression of appetitive state by pain. However, if reward experience is quite high, then a pain stimulus can evoke an approach turn in the forager by relieving suppression of Appetitive State by reward experience. The model predicts that with positive reward experience, as from Drug consumption, an extremely hungry animal may attack severely painful stimuli.

ASIMOV Simulation

Quantitative results from ASIMOV are obtained by controls on the interface console (Supplementary Methods). These can set prey and Drug populations, variables of satiation and reward experience, and apply pain in controlled settings.

Prey and drug selectivity was examined under "Drug-Free" and "Addicted" states. Specifically, for the Drug-Free state, associative strengths for Flab and Hermi were adjusted to maximums of 1 by pre-feeding the forager 15 of each prey. For each trial, Fixation of Variables was used to set Satiation and Reward Experience to specific values, and then, using Presentation Mode, Flab, Herm, and Drug were separately presented to the forager to test whether the forager made an appetitive or aversive turn. Satiation was set at values ranging from 0.01 to 1.0 and Reward Experience was set at values ranging from -20 to 20. For the Addicted state, the procedures were similar, except that associative strength for Drug was also set at a maximum of 1.

In Addiction Cycle Mode, the user observes the forager in different phases of the addiction processes, where availability of Drug changes over time, starting with only prey and no Drug, and then adding and removing the Drug. In the last phase, Drug is present with its odor signature, but does not provide any reward to test the effects of learning extinction.

ASIMOV is implemented in the graphic, agent-based programming language, NetLogo (Wilensky, 1999), and is available at https://github.com/Entience/ASIMOV.

RESULTS:

Effects of Satiation and Reward Experience on Prey and Drug Selectivity

ASIMOV's simple aesthetic sense is modulated by reward experience and satiation. **Figure 2.2** shows the effects of satiation and reward experience on the forager's selectivity for prey and Drug, under both Drug-Free state and Addicted states (Methods). Notably, under both states, high positive reward experience and high satiation both suppress preference-seeking behavior, as the forager avoids all prey and Drug at the highest levels. This satisfaction of preferences ends preference-seeking behavior. In the context of addiction, withdrawal manifests as negative reward experience, which only affects appetitive state and is a direct consequence of homeostatic plasticity in reward circuitry. The immediate effect of Drug consumption is positive reward experience. Consumption of either prey, noxious Flab or benign Hermi, increases satiation and provides relatively small negative or positive rewards, causing the forager to maintain a "normal" range of reward experience; whereas the Drug causes significant fluctuation in reward experience, as upon consumption it causes immediate extreme increase in positive reward experience, and over time can lead to negative reward experience due to desensitization. In the Drug-free state, where the forager is not exposed to Drug but learns associations of benign and noxious prey, low satiation leads to decreased selectivity for prey, while higher satiation leads to greater selectivity for benign prey, Hermi. As only prey are available for consumption in a Drug-free state, the forager maintains a normal range of reward experience.

If the forager learns all associations for prey and Drug (**Figure 2.2**, Addicted), selectivity between prey is similar to the Drug-free state. However, since Drug is valued more than all prey, Drug selectivity is enhanced even at high levels of satiation. During negative reward experience, like withdrawal, non-selective consumption of prey is increased, leading to increased satiation and in turn greater selectivity towards Drug.

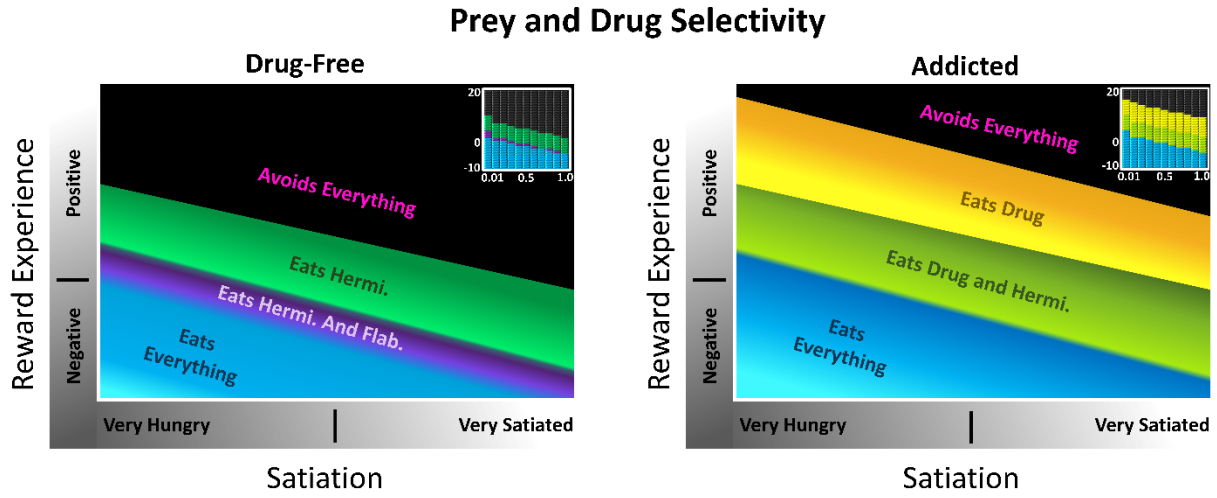

Figure 2.2. Effects of satiation and reward experience on selectivity for prey and Drug. Relations are smoothed from coarser quantitative data (insets), where satiation is varied from 0.01 to 1.0, and reward experience is varied from -10 to 20 in Presentation Mode. With enough available Drug in the environment, Drug consumption is favored over prey, leading to a lower nutritional state and low satiation. Selectivity is observed as an approach turn towards specific prey or the Drug. Left: A selectivity map for ASIMOV's forager naïve to the Drug. Learned associations for benign Hermi and noxious Flab are at maximum associative strength in this environment, with no learned association for the Drug. As prey consumption provides relatively small positive and negative reward experience, reward experience level is largely near zero. Right: An approach turn selectivity map for a forager addicted to Drug. Learned associations for all prey and Drug are at maximum associative strength. Drug consumption gives immediate positive reward and can eventually lead to negative reward experience during withdrawal; thus the forager's reward experience ranges from negative to positive. In negative reward experiences, like withdrawal, the effect of hunger is increased, and the forager shows less selectivity for prey and Drug. In high positive reward experience, there is increased selectivity for the Drug, so nutritional needs are often ignored in favor of Drug consumption.

With high positive reward experience, the effect of satiation on appetitive state increases, thereby enhancing selectivity for Drug. Thus, nutritional needs are ignored in favor of Drug

consumption. When satiation and reward experience are high enough, the forager becomes averse to all prey and Drugs until either state drops to a lower, permissive value. **Figure 2.2** thus examines the forager's dynamic aesthetic sense, showing shifts in preferences across differing satiation and reward experience, as well how these preferences change after a new experience.

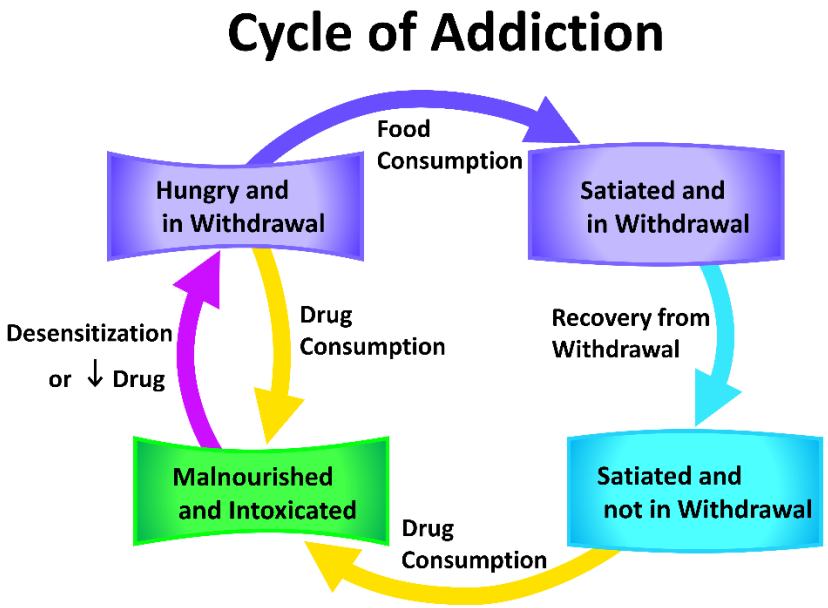

Figure 2.3. Four general states of ASIMOV's forager in addiction. With enough food and learned associations for prey and Drug, the cycle leads to Drug seeking, consumption, and poor nutritional state. In withdrawal, hunger has a stronger effect on appetitive state, reducing selectivity for Drug and prey consumption. In this dual state of hunger and withdrawal, the forager can consume the Drug non-selectively, which, if easily available, leads to intoxication and malnourishment. Without Drug it can consume prey non-selectively, increasing satiation. As the forager becomes satiated and recovers from withdrawal, satiation increases its selectivity for the Drug. If enough Drug is available and consumed, the forager's high positive reward experience reduces the effects of hunger to leave it in a state of malnourishment and intoxication. In this high reward experience, selectivity for the Drug is still increased (see **Figure 2.2**, right), and if sufficient Drug is available, consumption continues to maintain high reward experience. If insufficient Drug is available, or if the forager is too desensitized to Drug reward, it falls into withdrawal, begins feeling the effects of hunger more acutely, and starts the behavioral cycle over again.

The more specific and severe instance of the aesthetic process in addiction was further explored. **Figure 2.3** depicts the ASIMOV forager's generalized states in the Addicted state,

where if enough food is available, it enters a cycle of Drug-seeking behavior. Since high satiation increases selectivity for Drug, and withdrawal does not deter Drug consumption, without intervening circumstances the forager inevitably seeks out the Drug when its signature odor is present. So, available Drug naturally leads to high Drug consumption rates and lower nutritional state. Conversely, inadequate Drug supply leads to withdrawal and a period of overconsumption of prey. More significantly, however, the graphs of **Figure 2.2** indicate that the forager's strong preference for Drug is largely independent of fluctuations in reward experience and satiation. In the Addicted state the forager approaches Drug at most levels of satiation and reward experience, and its selectivity for Drug is higher overall than for any prey type. Thus, the primary driver for Drug consumption is the high associative strength and the resulting strong selectivity for Drug.

Phases of Addiction in a Dynamic Environment

The dynamics of the ASIMOV forager's aesthetic sense and its effect on foraging were explored further in a changing environment, without fixation of variables, using the Addiction Cycle Mode (**Figure 2.4**; Methods, Supplementary Methods). Here, the environment starts with 3 Flab and 3 Hermi, which respawn in random sites when eaten. Reward experience begins as positive as the forager learns associations for Hermi and Flab, then declines by homeostatic mechanisms. In the second phase, 6 Drug items are introduced. The forager finds and consumes the Drug, typically by accident when attempting to consume nearby prey. After first encounters, Drug consumption rate rises quickly. This coincides with a slight initial decrease in Hermi consumption, likely due to the decreased effects of hunger on the forager's appetitive state and

competition with Drug. As frequent drug consumption continues, desensitization to Drug reward increases, with marked fluctuations and decreasing average reward experience.

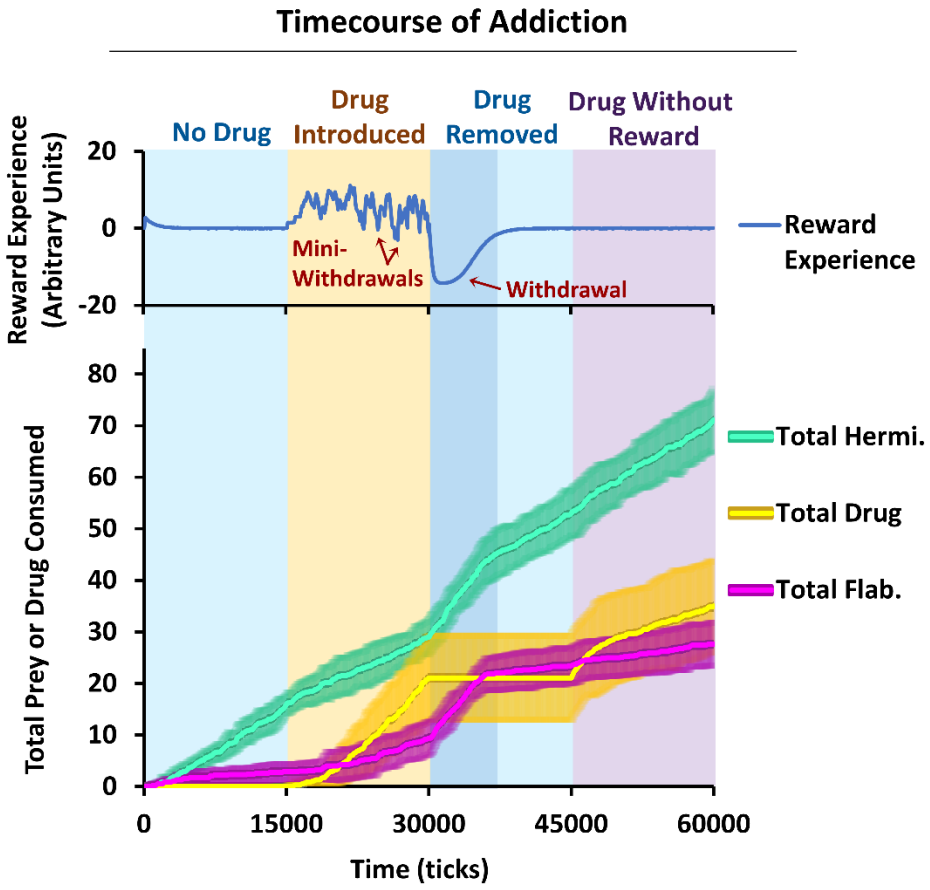

Figure 2.4. Phases of addiction. In Addiction Cycle Mode, Drug is introduced and removed. In these phases, the forager experiences desensitization, withdrawal and cravings. The curves displayed are timecourses of reward experience and the total numbers of different prey or Drug consumed, averaged over 10 trials of the Addiction Cycle Mode simulation. Error bars (SEM) are shown for total Hermi, Drug, and Flab consumed. When Drug is introduced, first-time Drug consumption typically occurs by accident or at low satiation. As Drug consumption continues, the forager experiences mini-withdrawals, wherein reward experience swings rapidly between positive and negative values. When Drug is removed, the forager undergoes withdrawal, during which loss of reward exacerbates the effects of hunger, increasing non-selective consumption of prey. In the Drug Without Reward phase, when Drug is first reintroduced with the same odor signature but no reward, the forager quickly resumes Drug consumption, even after recovery from withdrawal, since the associative strength for the Drug ("cravings") is still high. In later stages of Drug Without Reward phase, Drug consumption is significantly reduced due to the decrease in Drug's associative strength.

Thus, in an extended period of drug consumption, especially if Drug is limited, the forager begins to experience “mini-withdrawals”, where rapid switches occur between positive reward experience and brief periods of negative reward experience (**Figure 2.4**). Mini-withdrawals are caused by the forager’s inability to consume Drug quickly enough to overcome the marked desensitization developed in its HRC circuit, causing a brief negative reward experience. In this phase, the forager can markedly change its eating habits, oscillating from overeating prey in withdrawal to low nutrition during high reward experiences.

With removal of Drug, the forager enters withdrawal, followed by slow recovery. In withdrawal, consuming both Hermi and Flab increases markedly due to reduced prey selectivity. Notably, if Drug is reintroduced, consumption is resumed often more quickly than when Drug was first introduced, reflecting the effect of cravings for the drug. As Drug is reintroduced with no reward (**Figure 2.4**, Drug Without Reward), consumption quickly resumes at an initially high rate, which declines as the associative strength between Drug odor and reward decreases with extinction, approaching zero. This post-addiction phase is then similar to the initial Drug-free phase (**Figure 2.2**), and the Drug consumption rate becomes like that for a “neutral” prey without either nutrition or reward, and for which the forager has no associative strength. This suggests that the Drug consumption at the end of the phase is primarily due to accidental consumption and the effects of hunger, rather than a learned association.

Effects of Satiation and Reward Experience on Pain Threshold

Pain modifies the effect of reward experience on appetitive state and thus modulates the aesthetic sense. Since pain and reward experience are reciprocally inhibitory, a strong pain stimulus overrides the aversive effect of high reward experience to become the primary aversive

influence. This effectively alters the impact of reward experience. A positive reward experience can reduce the effect of pain and thus increases appetitive state, instead of reducing it as happens without pain. In contrast, negative reward experience aggravates the effect of pain.

To explore the approach-avoidance response to painful stimuli at different values of reward experience and satiation, in Presentation Mode ASIMOV's forager is immobilized but retains free turning responses, and pain stimuli are applied to the forager's right anterior region at a strength of 10. Reward experience affects pain thresholds for the aversive turns (**Figure 2.5**). Immediate rewards from Drug consumption effectively increase pain threshold, while withdrawal from Drug lowers it.

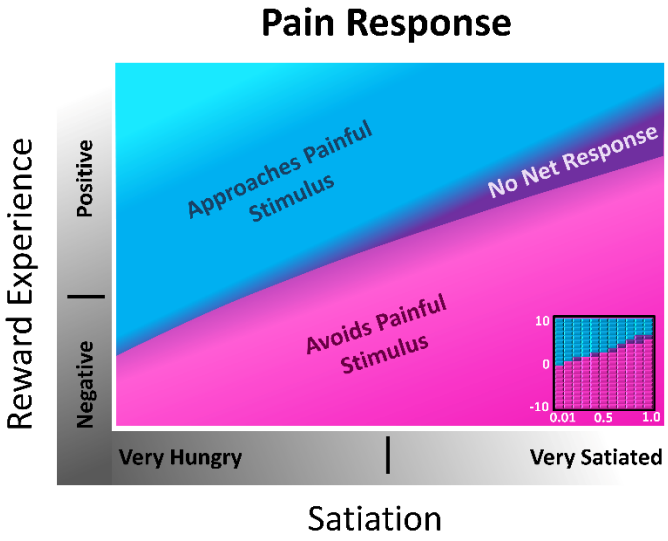

Figure 2.5. Effects of satiation and reward experience on response to pain. In general, positive reward reduces the ASIMOV forager's aversion to pain, while negative reward enhances it. Net responses to pain application are classified as approach turns to the stimulus, no net turn, or avoidance turns. Relations are smoothed from coarser quantitative data shown in the inset at the bottom right corner, where satiation is varied from 0.01 to 1.0, and reward experience is varied from -10 to 10 in the simulations. In a neutral reward experience, the forager makes an approach turn towards pain at very low satiation, and avoidance turns at higher satiation levels. A positive reward experience reduces the effect of pain, causing approach turns towards pain at higher satiation levels, while negative reward experience causes aversive turns at all levels of satiation.

At very low satiation and without reward, pain induces approach turns, but at a higher state of satiation pain causes avoidance. However, a positive reward experience immediately following Drug consumption reduces the effect of pain and causes approach turns even at even higher satiation levels. Negative reward experience, as occurs in withdrawal, worsens the effect of pain and causes aversive turns at all levels of satiation.

DISCUSSION:

ASIMOV reproduces a simple aesthetic sense, based on known neural circuitry of cost-benefit decision in foraging. Aesthetic valuation is a basic function in the ancient circuitry of foraging behavior, where generalist foragers establish preferences and aversions to the sensory signatures of different prey through reward experience. The aesthetic sense produces affective valuations that are expressed in behavior by characters of approach or avoidance. The signature stimuli connected with different prey acquire salience from interactions of reward learning and motivation, and confer ability to discriminate prey based on rewarding characters.

Reward learning allows opportunistic, foraging generalists that hunt in unpredictable environments to exploit prey available at different times and endowed with special qualities of nutrition or defense. Motivation acts with reward learning to organize cost-benefit analysis of predatory attempts, facilitating the negotiations of risk with need in foraging (Brown et al., 2018; Gillette et al., 2000). Reward learning likely has ancient origins, and is documented among generalist foragers in annelids, mollusks, insects, spiders, and even nematodes and flatworms. In parallel are aspects of the aesthetic sense in terms of abilities to evaluate stimuli, and specifically addictive behaviors (Barron et al., 2009; Carter et al., 2006; Entler et al., 2016; Heberlein et al.,

2009; Kusayama & Watanabe, 2000; Lee et al., 2008; Nathaniel et al., 2009; Palladini et al., 1996; Shipley et al., 2017).

The present results are consistent with addiction as an extreme expression of aesthetic choice. Strong learned association between Drug context (here odor signature) and high reward establishes preference for Drug at most values of satiation and reward experience (**Figure 2.2** and **Figure 2.3**).

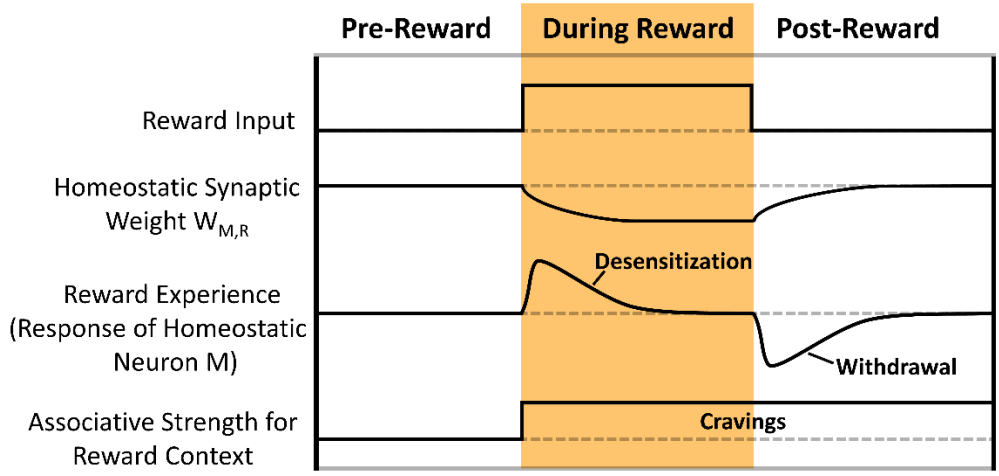

Figure 2.6. Effects of homeostatic plasticity in addiction. During reward input like Drug consumption, synaptic weight W changes dynamically based on the activity of neurons M and R. The response of neuron M to reward input desensitizes during prolonged reward input, and markedly decreases in withdrawal after loss of reward input. If reward input is paired with a stimulus such as Drug odor, after cessation of reward input the associative strength (incentive) for that stimulus remains high. This causes prolonged cravings for the high reward stimulus after completion of withdrawal.

The high-reward Drug stimulus causes strong association between reward context and actual reward. Lower reward stimuli, like the prey Hermi, cause milder effects because they produce a weaker learned association and less fluctuation in reward experience. The lesser effects are better seen as lower level "preferences", rather than addictions. Thus, while fluctuating reward experience acts together with learned context to influence Drug consumption in addiction, the major factors are learned association and homeostatic plasticity. For an addicted

animal, the model predicts that high drug availability leads to high drug consumption rates and lesser nutritional state, as is not unusual in actual drug addiction. In contrast, for a drug-naïve animal, drug consumption is less likely to occur under most circumstances but can be mediated either by chance or external agency (such as through peer pressure).

Corollaries of the addictive experience are desensitization to the rewarding properties of the addictive stimulus, withdrawal, a slow resensitization to drug reward, and prolonged cravings. Homeostatic plasticity, a use-dependent compensatory adjustment in the excitability of neurons and their networks (Turrigiano & Nelson, 2004), is strongly implicated in the addiction process, as when drugs affect action potential production or synaptic strengths. The "reward experience" is modified by homeostatic plasticity and thereby accounts for the dynamics of aesthetic valuations and characteristics of addiction. Homeostatic plasticity in ASIMOV's HRC module is responsible for desensitization and withdrawal by its use-dependent negative feedback to appetitive state. A likely HRC analog in the mammalian brain is the nucleus accumbens, which receives rewarding dopaminergic input, can suppress feeding via GABAergic projections to the appetite center in the lateral hypothalamus (Lutter & Nestler, 2009), and expresses notable homeostatic plasticity in the addiction process (Huang et al., 2011). While the negative affect of withdrawal is associated with reduced dopaminergic signaling in response to reward, it may also involve increased sensitivity of stress systems in extended amygdala, habenula, and hypothalamus (Volkow et al., 2019).

Figure 2.6 summarizes effects of reward learning and homeostatic plasticity in addiction. Desensitization induces more Drug seeking to keep reward effects high and to oppose the negative reward experience of withdrawal. In withdrawal, without reward the response of the reward circuitry (neuron M) induces negative affect. Then pain and hunger have greater impacts

on appetitive state. If cessation of reward input endures long enough, reward circuitry resensitizes and negative affect of withdrawal decreases. However, the association between reward context, such as Drug odor, and positive reward remains high, resulting in "cravings": marked increases in appetitive state and approach behavior whenever contextual stimuli (CSs) associated with the high reward (Drug), are encountered, analogous to powerful desire.

What is the adaptive significance of a homeostatically plastic reward circuit in an active forager in a natural environment? HRC, and its negative feedback to the feeding network, could normally function to maintain caloric intake in an environment with changing availability of different food sources. Thus, natural function may be to maintain caloric intake in "feast or famine" scenarios. When prey are readily available, rewarding food items could be common and consumed with high frequency. High frequency consumption would be maintained by desensitization to reward in the HRC, as inhibitory feedback to the feeding network would be reduced. In a famine scenario, if a preferred food source became scarce, ceasing its consumption would lead to some level of withdrawal, resulting in less HRC inhibition of the feeding network and a period of decreased selectivity for all food types, which could adaptively promote caloric intake.

In ASIMOV, desensitization to a repeated, moderately rewarding prey type causes only small withdrawal effects easily managed with consumption of other rewarding foods. Withdrawal from the high-reward Drug is more severe and the same high magnitude of reward would only be available from other addictive agents. Likewise, heroin users in withdrawal may resort to cocaine to alleviate withdrawal symptoms (Leri et al., 2003) or other opioid receptor agonists, such as methadone (Amato et al., 2004; Hirata & Castro-Alamancos, 2006). These agonists are also addictive, and withdrawal typically needs further treatment. More broadly,

some addictive behaviors have high rates of co-occurrence with substance use, such as gambling (Walther et al., 2012), reflecting common neurobiological and molecular pathways (Nestler, 2005). Notably, in ASIMOV withdrawal decreases selectivity among prey (**Figure 2.2** and **Figure 2.4**), potentially allowing the forager to seek novel stimuli that might bring some positive reward.

In ASIMOV's forager, satiation and reward experience modulate motivation in foraging and seeking behaviors. Both positive reward experience and high satiation suppress appetitive state, thereby promoting avoidance and higher selectivity in foraging. In contrast, negative reward experience and low satiation stimulate approach and lower selectivity. Similarly, the hypothalamic circuits governing hunger and satiety modulate the reward system, where hunger can increase the reinforcement, behavioral responsiveness, and seeking of drugs of abuse, while satiety signals generally reduce these effects (Cassidy & Tong, 2017; Zheng et al., 2012). Reward experience and pain have reciprocal effects. Negative reward experience in withdrawal exacerbates effects of both pain and hunger, while positive reward experience (like Drug consumption), can reduce effects of pain and hunger on appetitive state (**Figure 2.2** and **Figure 2.5**). The role of pain in the addiction process begins with the ability of rewarding stimuli and their learned cues to suppress its awareness (Altier & Stewart, 1999). This may be adaptive for the foraging animal when dealing with prey defenses (Gillette et al., 2000) or perhaps needing to ignore an injury to hunt. Hunger may also inhibit non-acute, inflammatory pain, though hunger itself may be suppressed by acute pain (Ponomarenko & Korotkova, 2018). Pain relief by itself can be rewarding (Navratilova & Porreca, 2014). Desensitization to drugs like cocaine and amphetamines acting on dopaminergic reward pathways can bring on painful side effects, perhaps in part because reward pathways that act to suppress effects of pain (Leknes & Tracey,

2008) are habituated. More seriously, withdrawal from opiates is worsened, as natural reward mechanisms are blunted and pain pathways simultaneously rebound from drug suppression with overshooting strength likely to also originate in homeostatic plastic mechanisms.

ASIMOV's forager developed a high rate of Drug consumption whenever it was available (**Figure 2.4**). First-time Drug consumption typically occurred when the forager was in a low satiation state, or by accident as when Drug was very close to nearby prey. When Drug was removed, withdrawal occurred, followed by recovery. When a new version of Drug was then introduced without reward, the forager immediately resumed a high Drug consumption rate, as it retained high associative strength for Drug, representing "cravings". But, as Drug was consumed without reward, the association for Drug extinguished, and consequently consumption rate fell significantly. These results reinforce the notion that associative strength with reward is a strong driver of addictive drug consumption. They also show that when the recurring context in which the Drug is acquired is paired with non-rewarding or aversive stimuli, it can diminish Drug consumption. This works well in simulation, where control of variables is rigid, but is not easily done in human populations.

Comparison to Other Models and Theories of Addiction

Basic characters of previous theoretical treatments emerge in ASIMOV's function. Notably, in the opponent-process theory of motivation (Solomon & Corbit, 1974), hedonic scale and the standard patterns of affective dynamics are analogous to the fluctuations in ASIMOV's reward experience before, during, and after reward input (**Figure 2.6**). Specifically, 1) in opponent-process the peak of the primary hedonic process theory corresponds to ASIMOV's reward experience in onset of rewarding input, 2) "hedonic adaptation" corresponds to

desensitization of reward experience, 3) “affective after-reaction” corresponds to withdrawal, and 4) “decay of after-reaction” corresponds to resensitization. Moreover, in the simple HRC module of ASIMOV, the primary hedonic process of opponent-process theory relates to neuron R’s response to direct reward input. The secondary opponent hedonic process, which is slow and initiated by the first process, is analogous to homeostatic plasticity of neuron M in the HRC (**Figure 2.1**, left). As in neuronal homeostatic plasticity, the secondary opponent process changes with use, such that rewarding effects are diminished and withdrawal effects increase. This is analogous to the cumulative effects of desensitization via the HRC.

Redish et al. (Redish et al., 2008) attributed the emergence of addiction to vulnerabilities in decision-making arising in animals’ systems for observation, planning, or habit. For ASIMOV, these correspond to sensory odor integration and the somatic mapping function, appetitive state (including HRC), and the reward learning algorithms, respectively. Two of the primary vulnerabilities obvious in ASIMOV are homeostatic dysregulation and overvaluation, which may have been among the first to emerge in the evolution of mechanisms underlying aesthetic and addictive processes.

Previous computational studies examined addiction through reinforcement learning (RL) models (Dezfouli et al., 2009; Redish, 2004), and RL actor-critic models (Takahashi et al., 2008). These did not account for homeostatic processes or the internal state of an organism. The Deperrois et al. (Deperrois et al., 2018) model of nicotine addiction used homeostatic down-regulation of receptors in addition to an RL framework; however, it did not account for internal state, and specifically focused on nicotine addiction’s effect on mammalian circuitry. Keramati et al. (2017) proposed an actor-critic model of homeostatic reinforcement learning (HRL) combining homeostatic and RL theories of addiction with the effects of organismal internal state,

suggesting that rewards calculated by the organism are naturally modulated by its internal state. ASIMOV also takes into account RL, homeostatic mechanisms, and internal state with much simpler calculations, in particular for internal state and its integration with learned associations and external stimuli.

Few computational models explore the origins of addiction or impulsivity from foraging circuitry. Barack and Platt (Barack & Platt, 2017) introduced a model of foraging decision comparing values of short-term options against long-term reward rates in iterated foreground-background accept-or-reject contexts, proposing that impulsivity, as seen in addiction, results from inaccurate estimation of long-term reward rates. In an area-restricted search (ARS) foraging model (Hills, 2006), Hills proposed that too much dopamine signaling was associated with a much too focused cognitive ARS, as may be the case in addiction. ASIMOV differs from these models in its architecture, but both the attentional aspect of ARS and the mis-estimation of long-term rewards resemble ASIMOV's forager over-learning and over-estimating Drug reward, and becoming extremely focused on attaining and consuming Drug, often ignoring other prey and hunger effects.

Berridge and collaborators described the emergence of addiction in the processes that incentivize rewarding stimuli (Berridge & Robinson, 2016). These authors differentiated "liking" and "wanting" in the addiction process. Liking describes "in-the-moment" hedonic experiences responding to stimuli, analogous to ASIMOV's immediate stimulus-driven change in reward experience. Wanting, the motivational drive that lends salience to incentive stimuli, is enhanced by learning of cues. In ASIMOV, wanting is embodied in the relations between the Appetitive State and Incentive modules that determine associative strengths to set the salience of incentivized stimuli (**Figure 2.1**).

Comparison to Other Models and Theories of Aesthetics

Aesthetics relies heavily on attention, as where objects considered beautiful or the opposite by observers will often draw their focus. Thus, aesthetics should also include a liking aspect, and the mere observation or experience of a beautiful object will bring pleasure often not associated with biological urgency (Mechner, 2018). As in addiction, the liking aspect of aesthetics must also depend on a reward system like the HRC, dissociated from physiological need and with internal dynamics independent from external reward input. For instance, the aesthetic pleasure received from observing a painting serves no immediate physiological purpose (Mechner, 2018). Boredom, habituation, and decreasing aesthetic pleasure received from a painting viewed multiple times may be explained by the homeostatic plasticity of a reward system like the HRC, where multiple encounters with the same pleasing stimulus, particularly over a short period of time, decrease the immediate change in reward experience and so decrease the aesthetic pleasure received.

Rolls (2011) explored an origin of aesthetics in goal-directed behavior, suggesting that gene-specified rewards and punishers can establish inherent aesthetic values. But the explicit, rational planning system also affects this valuation, allowing decisions made that might be in the subjective interest of the individual but not necessarily the genes. In ASIMOV, the genetically inheritable analogs include the forager's general attraction towards betaine, aversion to pain, suppression of appetitive state by satiation, and positive reward experience, and all other relations not included in its reward learning system; this in turn may be considered as the rational system. Both ASIMOV's genetic and rational systems, while fairly simple, are crucial to

decision-making in foraging, contributing to the final integration of appetitive state to specify the forager's goals for action (approach-avoidance), that may be aesthetically attractive or aversive. Xenakis and Arnellos (2014) proposed that aesthetic perception involves interactions where uncertainty is high, with no available relevant knowledge, and where emotions are used to evaluate interactive indications and thereby reduce uncertainty. This is most closely related to ASIMOV's reward experience, which becomes a major factor in setting appetitive state to make an approach or avoidance decision when the agent is in an uncertain, unfamiliar situation, where it cannot rely on learned or innate associations.

Dissanayake (2015) provides an interesting perspective on evolution of art and aesthetic appreciation from proto-aesthetic operations founded in adaptive ancestral mother-infant interactions. A mother's vocalizations with her infant, coinciding with increased parental care in human evolution, is an example of a "make special" primitive, where ordinary things are deliberately made significant, and is one of the important ingredients of art. ASIMOV might provide some insight for the origin of even simpler aesthetic primitives. A primitive that should precede a make special primitive, entails the existence of specific attention mechanisms. Indeed, reward experience and incentive, and how they modulate appetitive state, involve very simple attentional mechanisms, with homeostatic plasticity of the HRC potentially relating to attentional habituation and sensitization.

To our knowledge there are no other computational models exploring aesthetics in a foraging context. There are computational models of aesthetics and creativity, however these are mainly limited to the field of machine learning and information theory. Schmidhuber (Schmidhuber, 2010), for instance, introduced an intrinsically-motivated agent-based model, and proposed that creativity and associated behaviors could be driven by a simple algorithm using

reinforcement learning to maximize an agent's internal joy to discover and create novel patterns. While ASIMOV does not address novelty seeking, the homeostatic plasticity of the reward system can explain why a reward input that is given repeatedly loses its effect with its novelty. It thus becomes less aesthetically pleasing, revealing characters of boredom and providing a basis for seeking novelty. Thus, ASIMOV might be easily developed further, so that stimulus-specific reward experiences would decline, mimicking boredom and promoting seeking of new aesthetic experiences.

ASIMOV's Limitations

ASIMOV is a relatively simple model that does not take on all the intricate dynamics of aesthetics and addiction in humans and other mammals. ASIMOV's forager makes the simplest of decisions for approach or avoidance turns. There are no multi-step decision-making processes and no complex motor output. The Rescorla-Wagner algorithm for learning used is one of the simplest; it does not simulate episodic or sequenced memory and is less complex than the reinforcement learning algorithms employed in other models of addiction or aesthetics. ASIMOV's architecture is largely based on *Pleurobranchaea*'s circuitry for foraging decisions, with linear and sigmoidal relations between elements, rather than on mammalian learning, reward, and decision circuits modeled with spiking neurons. While this reduces biophysical realism for ASIMOV, it is significant that the minimal model captures common origins of addiction and aesthetics in foraging circuitry. Further expansions of ASIMOV for sequence learning and simple episodic memory may greatly enhance the forager's aesthetic.

CONCLUSION:

Addiction emerges as an extreme expression of aesthetic preference. The consequences of the addictive experience are desensitization to the rewarding properties of the addictive stimulus, withdrawal, a slow resensitization to drug reward, and prolonged cravings. Homeostatic plasticity, a use-dependent compensatory adjustment in the excitability of neurons and their networks, is strongly implicated in the addiction process. The modification of the reward experience by homeostatic plasticity thus accounts for the dynamics of both aesthetic valuations and characteristics of addiction. The relations prominent in the ASIMOV simulation – strong preferences, desensitization, withdrawal, resensitization, and protracted cravings – overlap with those that attend the highs and lows of social relationships (Burkett & Young, 2012) and compulsive behaviors like gambling, shopping, internet use, and self-harm (Blasco-Fontecilla et al., 2016; Rosenberg & Feder, 2014; Walther et al., 2012). Reward learning, as well as reward experience and its relationship to pain and modulation by homeostatic plasticity, are causally central to these conditions.

These relations may also lie at the root of innate and learned aesthetic preferences in food, music, and art, as well as the drive behind creative activities. The common relations suggest that the diverse aesthetic processes of affective valuation in higher vertebrate experience are evolutionary derivatives of the basic neuronal circuitry of foraging economics, put to different functions but conserving similarities in their overall organization. Evidence to test this hypothesis is presently scant, but might be usefully sought in comparative studies.
ASIMOV is an easily accessible agent-based simulation, where the decisions and movement of the forager are readily observable, and the interface allows for easy user interaction, including control of the forager's behavior, movement, and environment. The software is highly accessible

to verification and experiment, and is available on the internet (https://github.com/Entience/ASIMOV) for examination, use, and modification.

SUPPLEMENTARY METHODS:

The ASIMOV model is derived from the original Cyberslug foraging simulation (Brown et al., 2018) with essential modifications. Cyberslug presents the logic of foraging decision based on reward learning and motivation, and is available at https://github.com/Entience/Cyberslug. ASIMOV is implemented in the same graphic, agent-based programming language, NetLogo (Wilensky, 1999), and is available as extended data at https://github.com/Entience/ASIMOV. NetLogo software is chosen for its simple availability and accessibility to a broad audience. The ASIMOV forager encounters two virtual prey in its environment, the benign Hermi and noxious Flab, named after prey sea-slugs *Hermissenda crassicornis* and *Flabellina iodinea* prey that *Pleurobranchaea* encounters in the wild (Noboa & Gillette, 2013). Each prey secretes two odors: the resource signal *betaine*, a predictor of nutritional resource (Gillette et al., 2000), and either of "*odor_hermi*" or "*odor_flab*." Odors diffuse realistically over time and space. Prey move in simple random walks. Prey numbers are constant; when consumed, replacements appear at random positions. The specific odors of the prey Hermi and Flab become associated with positive and negative expected rewards, respectively, resulting in learning preference for Hermi and avoidance of Flab. These positive and negative associations are established in reward learning via the Rescorla-Wagner algorithm for classical conditioning (Rescorla & Wagner, 1972). These effects are analogous to the actual predator's learned preferences for specific prey: ready consumption of the beneficial *Hermissenda*, and the rejection and aversive learning for the noxious Flab (Noboa & Gillette, 2013). Also present is the analog of an addictive Drug, a high-

reward item with its own sensory signature, which provides no nutrition. If prey nor Drug no longer provide reward upon consumption, the corresponding learned association decreases in strength via an extinction mechanism of the Rescorla-Wagner algorithm, in which the received reward is set to zero.

Appetitive State

The core of the AIMOV forager model, appetitive state, controls the choice of an approach or avoidance turn, and is a final integration of an animal's motivational state with stimulus incentive and pain. Motivational state itself is taken as a composite of satiation, a function of nutrition, and feedback from reward experience. By default, when appetitive state is low, the response to stimuli is avoidance, as in actual *Pleurobranchaea* (Gillette et al., 2000). Increasing appetitive state inverts the turn response direction to one of approach. Thus, appetitive state determines sensory thresholds for the approach turn toward prey and subsequent feeding responses. When high enough, corollary outputs from the feeding network switch the excitatory sensory input of the stimulus from one side of the turn network to the other, resulting in a turn towards the stimulus. Appetitive State is expressed as:

$$AppState = 0.01 + \frac{1}{1 + e^{-S}} + c_5 \cdot (AppStateSwitch - 1) \ , \tag{2.1}$$

$$S = c_1 \cdot Incentive - c_2 \cdot Satiation - c_3 \cdot Pain - c_4 \cdot PainSwitch \cdot RewardExperience \, , \tag{2.2}$$

$$PainSwitch = 1 - \frac{2}{1 + e^{-c_6 \cdot (sns\_pain - 0.2)}} \ , \tag{2.3}$$

where c_1, c_2, …, c_6 are constants available in the code. AppStateSwitch acts as a threshold-based switch between avoidance and approach turning behavior, with (AppStateSwitch - 1) causing a transient suppression of AppState during avoidance turns.

Incentive mixes intrinsic positive and negative properties of stimuli to integrate with memories of previous reward. In absence of incentivized sensory input or pain input, appetitive state is simply the basal motivational state, which in feeding behavior is based on integrating satiation, pain, and reward experience feedback. With addition of incentive, appetitive state becomes equivalent to "incentive salience" as defined in mammals (Berridge & Robinson, 2016). Satiation determines the baseline excitation state of the feeding network. At very low satiation, appetitive state is high, resulting in an approach turn towards any odor, even if novel. Sensory inputs integrate with memory into incentive. Incentive sums with satiation in the feeding network to either increase or decrease appetitive state.

ASIMOV extends the original model of approach-avoidance decision (Brown et al., 2018) by adding reward experience (**Eq. (2.1)**), explicit pain sensation, and reciprocal inhibition between them via the PainSwitch variable.

Reward Experience

Reward experience, an animal's activation of reward circuitry under the influence of homeostatic mechanisms, is expressed as a function of the output of the HRC module (**Figure 2.1**, left, and **Eqs. (2.4)-(2.6)**):

$$Reward\ Experience = -c_7 + \frac{2c_7}{1 + e^{M_0 - M}}, \tag{2.4}$$

$$M = W \cdot R, \tag{2.5}$$

$$\frac{dW}{dt} = \frac{M_0 - M}{\tau \cdot R} \quad , \tag{2.6}$$

where c_7 and τ are constants available in the code. In the model (**Figure 2.1**, left), a high reward input, as from a recreational drug, is amplified by neuron R. The amplified reward then feeds to a postsynaptic neuron M, whose capacity for homeostatic plasticity is analogous to habituation (McSweeney & Murphy, 2009). Specifically, the synaptic weight W between presynaptic neuron R and postsynaptic neuron M changes dynamically based on both presynaptic and postsynaptic activity, as well as on baseline activity, M_0, of neuron M (**Eq. (2.6)**). Neuron M activity is computed as the product of its synaptic weight W and the activity of neuron R (**Eq. (2.5)**), and fed to the Feeding Network. With repetition of a large rewarding stimulus or a long enduring Drug reward, the reward response of neuron M desensitizes by homeostatic plasticity. Notably, desensitization reduces positive reward effects (such as Drug reward), which also decay faster. Cessation of strong reward causes severe withdrawal. The magnitude of repeated rewards is proportionate to the rate of desensitization and magnitude of withdrawal.

Pain

Pain modifies the effect that reward experience has on appetitive state (**Eq. (2.2)**). Since pain and reward experience are reciprocally inhibitory processes (**Figure 2.1**, left), strong pain stimuli override the general aversive effect of high reward experience to become the primary aversive influence. This reciprocal inhibition is performed via the Pain Switch variable defined in **Eq. (2.3)**, which produces a sign change from +1 to -1, when pain is high. This effectively alters the effect of reward experience on appetitive state. A positive reward experience can actually reduce the effect of pain and thus promote appetitive state, instead of inhibiting it as it does normally. In contrast, a negative reward experience aggravates the effect of pain.

ASIMOV Interface

Quantitative results from ASIMOV are obtained by the controls on the interface console (Fig. S1). Important controls on the left side are Prey and Drug Population Controls, Fixation of Variables, Pain Application Controls, and Presentation Mode, and on the right are Addiction Cycle Mode. Prey and Drug Population Controls let the user control the amount of Hermis, Flabs, and Drug present in the environment. Fixation of Variables lets the user fix values for satiation, reward experience, and incentive. Satiation can range from 0.01 to 1.0, reward experience to -20 to 20, and incentive to -10 to 10. The default ranges are adjustable by editing the interface in NetLogo. Pain Application Controls allow the application of a painful stimulus to ASIMOV's forager. The strength of the applied painful stimulus can be adjusted by the Apply_Pain slider, ranging from 0 to 30 arbitrary units. The Poke-Left and Poke-Right buttons apply pain to the anterior left or right side part of the forager, respectively.

Presentation Mode controls the forager's prey and Drug intake, and tests approach and avoidance responses to specific prey or Drug. The forager can be immobilized except for approach-avoidance turns, and can be force-fed or presented with Hermi, Flab, or Drug to adjust the corresponding associative strengths, or to monitor the effects on reward experience. Presentation with a stationary Hermi, Flab, or Drug near its left side tests its turning response in terms of approach or avoidance towards the prey or Drug.

Addiction Cycle Mode, when enabled, allows the user to observe ASIMOV's forager as it freely forages and experiences different phases of the addiction processes of desensitization, withdrawal, and cravings. In this mode the availability of Drug changes over time, starting with an environment with only prey and no Drug, and then adding and removing the Drug, causing

the forager to go through desensitization, withdrawal, and cravings. In the last phase, the Drug is present with the Drug odor signature, but does not provide any reward on consumption. In this phase, Drug consumption decreases significantly (see Fig. 6 for results from the Addiction Cycle Mode). The Addiction Cycle Mode lasts 60000 software cycles (ticks), in which each phase lasts 15000 ticks. During an initial "No Drug" phase, the environment contains only the prey Hermi and/or Flab, letting ASIMOV's forager learn the corresponding associations. In the second "Drug Introduced" phase, the Drug is introduced for the first time with a high reward on consumption. The third phase is "Drug Removed" and the fourth and last phase is "Drug without Reward", where the Drug is reintroduced with the same odor signature, but provides no reward on consumption, causing ASIMOV's forager to decrease its associative strength for the Drug via a Rescorla-Wagner algorithm for extinction.

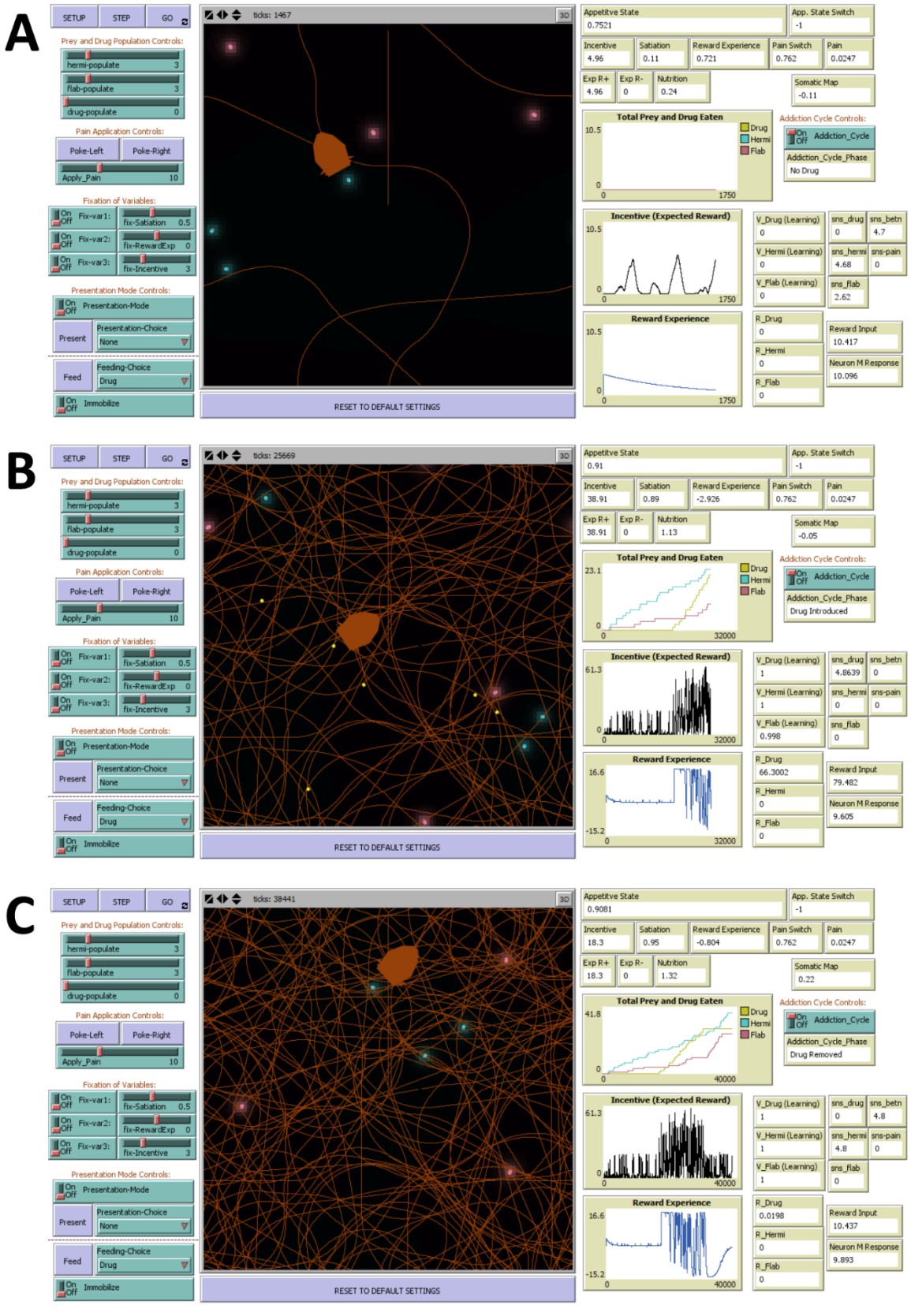
A
SETUP
STEP
GO
Prey and Drug Population Controls:
hermi-populate 3
flab-populate 3
drug-populate 0
Pain Application Controls:
Poke-Left
Poke-Right
Apply_Pain 10
Fixation of Variables:
Fix-var1:
fix-Satiation 0.5
Fix-var2:
fix-RewardExp 0
Fix-var3:
fix-Incentive 3
Presentation Mode Controls:
Presentation-Mode
Present
Presentation-Choice
None
Feed
Feeding-Choice
Drug
Immobilize
ticks: 1467
3D
RESET TO DEFAULT SETTINGS
Appetitve State
0.7521
App. State Switch
-1
Incentive
4.96
Satiation
0.11
Reward Experience
0.721
Pain Switch
0.762
Pain
0.0247
Exp R+
4.96
Exp R-
0
Nutrition
0.24
Somatic Map
-0.11
Total Prey and Drug Eaten
Drug
Hermi
Flab
Addiction Cycle Controls:
Addiction_Cycle
Addiction_Cycle_Phase
No Drug
Incentive (Expected Reward)
Reward Experience
V_Drug (Learning)
0
V_Hermi (Learning)
0
V_Flab (Learning)
0
sns_drug
0
sns_betn
4.7
sns_hermi
4.68
sns-pain
0
sns_flab
2.62
R_Drug
0
R_Hermi
0
R_Flab
0
Reward Input
10.417
Neuron M Response
10.096
B
ticks: 25669
Appetitve State
0.91
Incentive
38.91
Satiation
0.89
Reward Experience
-2.926
Pain Switch
0.762
Pain
0.0247
Exp R+
38.91
Exp R-
0
Nutrition
1.13
Somatic Map
-0.05
Addiction_Cycle_Phase
Drug Introduced
V_Drug (Learning)
1
V_Hermi (Learning)
1
V_Flab (Learning)
0.998
sns_drug
4.8639
sns_betn
0
sns_hermi
0
sns-pain
0
sns_flab
0
R_Drug
66.3002
R_Hermi
0
R_Flab
0
Reward Input
79.482
Neuron M Response
9.605
C
ticks: 38441
Appetitve State
0.9081
Incentive
18.3
Satiation
0.95
Reward Experience
-0.804
Pain Switch
0.762
Pain
0.0247
Exp R+
18.3
Exp R-
0
Nutrition
1.32
Somatic Map
0.22
Addiction_Cycle_Phase
Drug Removed
V_Drug (Learning)
1
V_Hermi (Learning)
1
V_Flab (Learning)
1
sns_drug
0
sns_betn
4.8
sns_hermi
4.8
sns-pain
0
sns_flab
0
R_Drug
0.0198
R_Hermi
0
R_Flab
0
Reward Input
10.437
Neuron M Response
9.893

Figure 2.7

Figure 2.7. Screenshots of the ASIMOV environment and interface. There are three user modes. In the default foraging mode, the forager (orange) encounters Hermi (blue orbs), Flab (pink orbs), and Drug (yellow orbs) and traces its path (orange contours). The user can select numbers of prey and Drug items in the environment and fix values of Satiation, Reward Experience, and Incentive. In Presentation Mode, the forager is immobilized and its prey and Drug intake are controlled to monitor approach and avoidance responses. Pain stimuli of adjustable magnitude can be applied to the left or right side of the ASIMOV agent's head to test the turn response. On the right of the interface, Addiction Cycle Mode, when enabled, allows the forager to forage freely and experience different phases of addiction in a dynamic environment. Drug availability changes over time, causing the forager to experience addiction, desensitization, withdrawal, and cravings (see text). **A)** ASIMOV's agent forages in the No Drug phase of the Addiction Cycle. Note that it initially performed several avoidance turns, and at the moment of the screenshot it is making an approach turn towards a prey. **B)** ASIMOV's agent forages in the Drug Introduced phase of the Addiction Cycle. Drug consumption significantly increases the forager's reward experience. As ASIMOV's agent continues Drug consumption, there is desensitization to the reward received from the Drug, and thus more fluctuation in reward experience, as seen in C. **C)** In the Drug Removed phase of the Addiction Cycle, without Drug access ASIMOV's forager undergoes withdrawal, represented in negative reward experience.

ACKNOWLEDGEMENTS AND CONTRIBUTIONS:

Marianne Catanho (MC) produced an early working model of addiction incorporating homeostatic plasticity. Ekaterina D. Gribkova (EDG) reworked, elaborated, and exhaustively tested the model with the HRC module and pain. EDG and Rhanor Gillette (RG) wrote the paper. This study was supported by Navy N00014-19-1-2373. Early stages of this work were supported by National Science Foundation grant IOB 04-47358 and National Institutes of Health grant R21 DA023445.

CHAPTER 3: LEARNING AND MEMORY

Learning and memory are the processes of encoding, storage and retrieval of information about past experiences, and in living organisms they are crucial for generating predictions that may aid in the organism's survival. Notably, learning and memory systems can enhance behavioral complexity in AI agents and animals alike. There have been many different models of memory and learning, widely ranging in complexity from simple ANNs with Hebbian learning, to full-scale modeling of hippocampal neurons using multi-compartment Hodgkin-Huxley models. AI applications, including virtual and autonomous agents, typically use simpler learning rules and network architectures that often do not resemble the biological substrates of learning and memory. In this chapter we review commonly used learning rules for ANNs and introduce two models of memory that operate on different scales. The first, the Feature Association Matrix (FAM), uses architecture similar to that of the auto-associative network in CA3 of hippocampus to memorize simple sequences, and the second model, Synaptic Input Time Difference Learning (SITDL), is a model of how neuronal synapses may be able to learn and memorize the time difference of input signals through changes in post-synaptic receptor populations.

ANN Learning Rules and Sequence Learning

ANNs with supervised, unsupervised, and reinforcement learning have been proposed for memory function and related brain circuitry. Supervised learning rules necessarily involve labeled data, while in unsupervised learning the network must learn something about the dataset provided, such as a hidden structure or how to organize it, without any given labels (Shalev-Shwartz & Ben-David, 2014). Note that in many ANNs with supervised learning, there appears to be a lack of biological correlate for the backpropagation algorithms used. A common

difficulty with unsupervised learning is that there may be no clear success evaluation procedure as there is for supervised learning. Reinforcement learning, on the other hand, involves optimization of decision processes in a goal-directed agent, where actions determine transitions between states, which may provide reward. Reinforcement learning, unlike unsupervised learning, does not attempt to find hidden structure in a dataset, but rather it attempts to maximize a reward signal through exploration and exploitation of action strategies (Sutton & Barto, 2018). Notably, there have even been models that use all three types of learning rules for brain-like learning and memory systems, where computational analogs of the cerebellum, cerebral cortex, and basal ganglia, are responsible for supervised, unsupervised, and reinforcement learning, respectively (Sasakawa et al., 2008).

Unsupervised Learning Rules

One of the simplest unsupervised learning rules for ANNs is the Hebbian learning rule (Yu et al., 2017), where the synaptic weight (w_{ij}), between neurons j and i, is updated as follows:

$$\Delta w_{ij} = A x_i x_j \tag{3.1}$$

with A being the learning rate, x_j typically being the pre-synaptic neuron's activity, and x_i as the post-synaptic neuron's activity, and where a neuron's activity is often simply defined as the synaptically weighted sum of all of its inputs. Note that the most basic Hebbian learning rules in ANNs primarily deal with rate-coding; in particular, changing synaptic weights according to pre- and post-synaptic activity rates. Therefore, learning in ANNs with temporal coding must employ temporal learning rules, such as spike-timing-dependent-plasticity (STDP), tempotron rules, SpikeProp rule, SPAN rule, Chronotron rule, and ReSuMe rule. STDP can be used with unsupervised learning, while most of the other aforementioned learning rules are for supervised

learning. Unlike the basic Hebbian rule, STDP takes into account the time difference (Δt) between the pre-synaptic and post-synaptic spikes. Here, the change in synaptic weight (Δw_{ij}) is characterized by **Eq. (3.2)** (Yu et al., 2017):

$$\Delta w_{ij} = \begin{cases} A^{+} \cdot e^{\left(\frac{\Delta t}{\tau^{+}}\right)}, & if\ \Delta t \leq 0 \\ A^{-} \cdot e^{\left(\frac{\Delta t}{\tau^{-}}\right)}, & if\ \Delta t > 0 \end{cases} \tag{3.2}$$

where A^{+} and A^{-} denote the learning rates and τ^{+} and τ^{-} denote the time constants. With STDP, if the pre-synaptic neuron fires before the post-synaptic neuron, the synaptic weight will grow, and if the pre-synaptic neuron fires after the post-synaptic neuron, the synaptic weight will decrease. Note that the vast majority of unsupervised learning rules for ANNs employ changes in synaptic weight, and there has not been an in-depth exploration of learning rules that employ changes in the timing of synaptic inputs (see **Learning Precise Timing through NMDAR Conductance Mismatch in a Computational Model of Developmental Synaptic Plasticity**).

Sequence Learning with Recurrent ANNs

Traditional feedforward ANNs typically cannot encode sequences of inputs. An issue is that the set of inputs presented to a neural network at some time point t, can be absent when a new set of inputs is presented at time t+1. Therefore, the neural network itself must somehow keep a memory trace of the input set at the previous timepoint in order to associate it with the next input set. Some solutions to this are to modify the structure of the ANN, or the learning rules, in order to keep temporary memory traces of the input sets that can then be used for forming associations.

In particular, recurrent neural networks (RNNs) take sequences as an input (Lipton et al., 2015), which can include temporal sequences. The primary difference between feedforward ANNs and RNNs is that the edges of RNNs span adjacent time steps (**Figure 3.1**), called recurrent edges, which essentially allows it to form associations between inputs introduced at different times. Long short-term memory (LSTM) models are special types of recurrent networks with memory cells that can learn even longer and more complex sequences and can overcome the problem of vanishing gradients that normal RNNs cannot. Furthermore additional auto-associative structure has been utilized to improve existing RNNs as well (Zhang & Zhou, 2017).

It seems few models explore learning rules and other architectures that might enable sequence learning, and reduce the need for typical RNNs, which appear to have no direct analogue in biological systems. In the next section, we explore learning and memory in biological systems, focusing on the brain architectures of mammalian memory, and introduce a model, FAM, that uses similar architecture and specialized learning rules for sequence memorization.

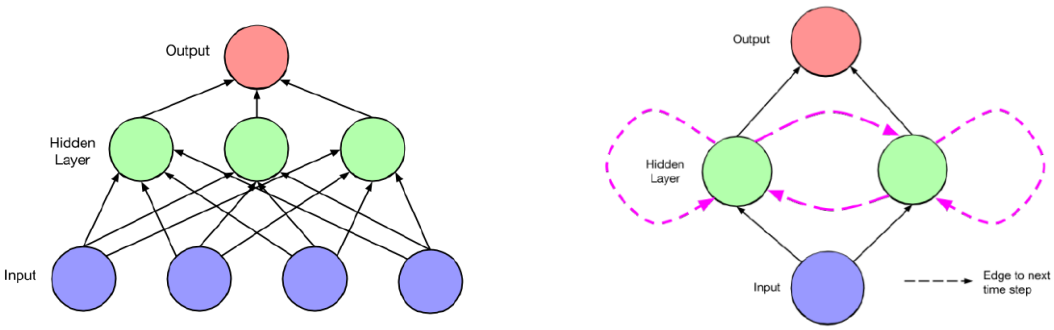

Figure 3.1. Feedforward neural network versus a recurrent neural network (right). Feedforward networks cannot learn sequences, whereas recurrent neural networks can, due to their edges spanning time steps. Figures reproduced from Lipton et al. (2015).

Feature Association Matrix for Simple Episodic Memory Enabling Sequence Learning

INTRODUCTION:

In mammalian brains, there are multiple types of long-term memory systems. Semantic memory, for instance, is for encoding general knowledge of the world, including the meanings of words, without any additional contexts like time or place. Episodic memory, also commonly known as autobiographical memory, is a memory for past events and experiences that, unlike most other types of memory, specifically encodes context (Gershman & Daw, 2017; D. M. Smith & Mizumori, 2006; Subagdja & Tan, 2015). Because episodic memory takes into account the temporal context of experiences, it essentially lets animals learn and recall sequences of events. While semantic memory consists of learning an association between sensory input and reward that is unbound in time and relation to other events, with episodic memory, a sequence of features, comprising events and physical stimuli, can be learned by association with specific rewards (M. W. Jung et al., 2018; D. M. Smith & Mizumori, 2006). Semantic and episodic memories that are not associated with a positive or negative reward, or any significant emotional state, are more likely to be forgotten. In an experimental example, semantic memory would allow a mouse to generally associate a lever press with receiving a reward, typically in the form of sugar water or food pellets, without encoding any other additional context. On the other hand, episodic memory can enable the mouse to learn the specific sequence of turns it needs to take in a maze in order to reach the end and receive a reward. In this case, each memorized turn has its own context and place within that sequence. If the reward at the end of the maze is removed, the memorized sequence would typically no longer hold any value for the mouse, and would typically be extinguished.

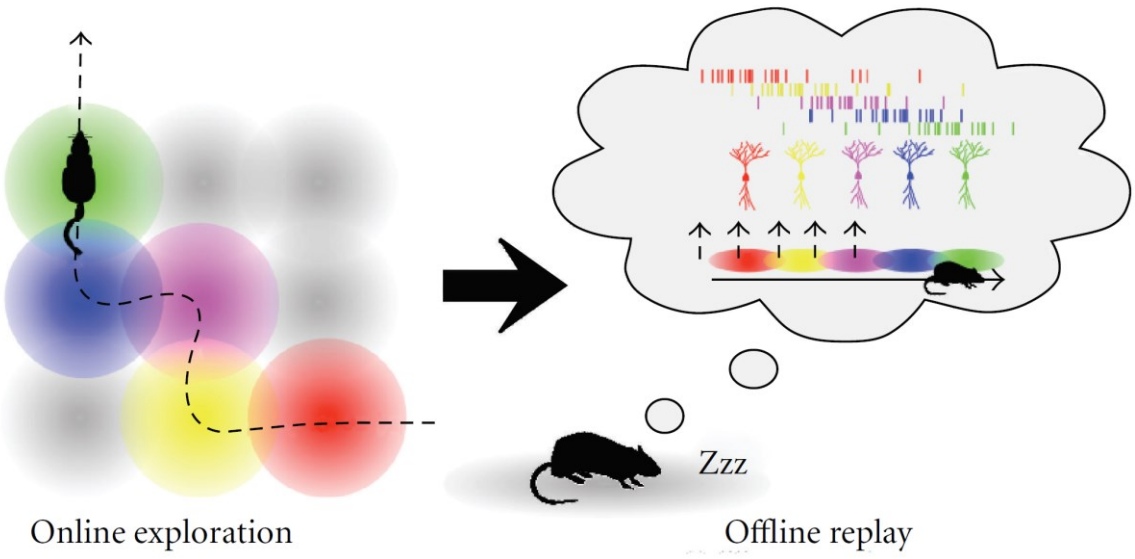

Figure 3.2. Each hippocampal place cell (right bubble, neurons), visualized in a different color, fires preferentially at a specific location as the mouse runs along a specific spatial path (left). This sequence of place cell activation is formed as the mouse runs the pictured spatial path, and the sequence can be replayed during sleep (right). Figure reproduced from Sadowski et al. (2011).

In mammals, episodic memory formation is mediated by the hippocampus, a specialized brain structure which contains a variety of network architectures and plasticity mechanisms. In the hippocampus, the firing of specific "place" cells represents specific spatial and temporal context for episodic memories (D. M. Smith & Mizumori, 2006). As a mouse runs through the maze, for instance, each place cell will preferentially fire at a specific point along the maze, and a sequence of place cell firing is formed. **Figure 3.2** shows that as a mouse runs a certain spatial path (**Figure 3.2**, left), specific neurons in the hippocampus will fire spike trains at specific points along that path (**Figure 3.2**, right bubble). This firing pattern can be replayed during sleep (**Figure 3.2**, right), which is thought to play a role in consolidating the memory of that spatial path. If the mouse has run the maze through several trials, all of which were rewarded, its hippocampus will exhibit a "forward replay" of the place cell sequence as it moves. If it pauses

along the maze, notably, there will also be "reverse replays" of the place cell sequences, as if the mouse is remembering how to proceed or what turn to make next to eventually reach the reward.

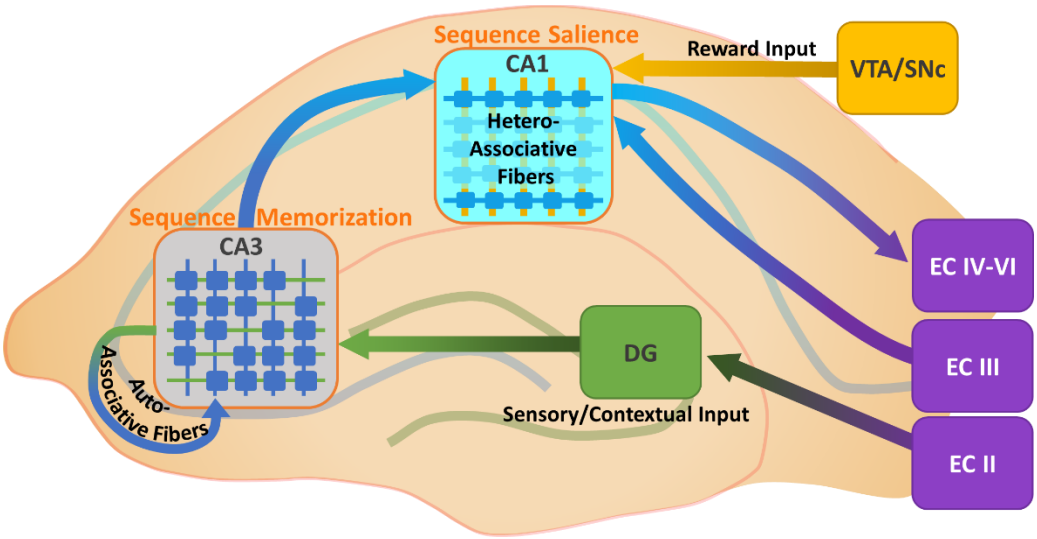

Figure 3.3. A hypothesis of sequence memorization in the hippocampus. The image presented here is that of the hippocampus, with substructures shown as DG, CA3, and CA1. CA3 sustains sequence formation and memorization, receiving sensory and contextual input from dentate gyrus (DG) and layer II of the entorhinal cortex (EC), as well as input from auto-associative fibers. Place cells in CA3 can be activated in sequence. This sequence of activation can be "memorized" through appropriate changes in synaptic strengths and replayed in forward or reverse play. CA1 also has its own place cells, and it receives input from CA3 and reward input from dopaminergic midbrain structures such as the ventral tegmental area (VTA) or substantia nigra compacta (SNc), forming a network of hetero-associative fibers. CA1 is a likely site for assigning sequence salience and modification of sequence stability, where sequences of place cell activations can be paired with reward input. Figure is partially adapted from Wilson et al. (2006).

Figure 3.3 gives a simplified overview of the hippocampal circuits that learn sequences of sensory cues, such as those responsible for the place field coding. The entorhinal cortex (EC) is a multi-layered brain structure that is the main gateway between the hippocampus and the rest of the brain. Layer 2 of the entorhinal cortex (EC), in particular, provides sensory and contextual input to the dentate gyrus (DG) of the hippocampus, where further processing takes place, such as pattern separation of similar input patterns. This processed sensory and contextual input from

DG activates specific neurons, which are the place cells, in CA3. Each place cell in CA3 is associated with a specific location or context for the animal. CA3 also has an auto-associative network, meaning that each place cell synapses onto, or outputs to every other place cell. These auto-associative synapses have been shown to use spike-timing-dependent plasticity (STDP), and are essentially responsible for the learning and formation of the place cell navigation sequences (Hasselmo et al., 1995; Van Strien et al., 2009). CA3 outputs to CA1 as well, which has its own place cells that are thought to be more sensitive to reward input. In particular, the assignment of reward value to sequences formed in CA3 may occur through CA1 of the hippocampus (M. W. Jung et al., 2018), which receives dopaminergic input from midbrain substantia nigra compacta and ventral tegmental area. **Figure 3.4** shows the difference in neuronal firing at CA1 and CA3 of the hippocampus when a mouse traverses a rewarded path versus a non-rewarded path. Clearly, CA3 does not differ much in the intensity of place cell firing, while CA1 does, with place cell activation being greater during the high-value sequence than during the low-value sequence. Notably, the processing of reward, as encoding of reward magnitude or reward prediction error, is uniquely associated with the reverse replay of sequences, rather than forward replay (Ambrose et al., 2016).

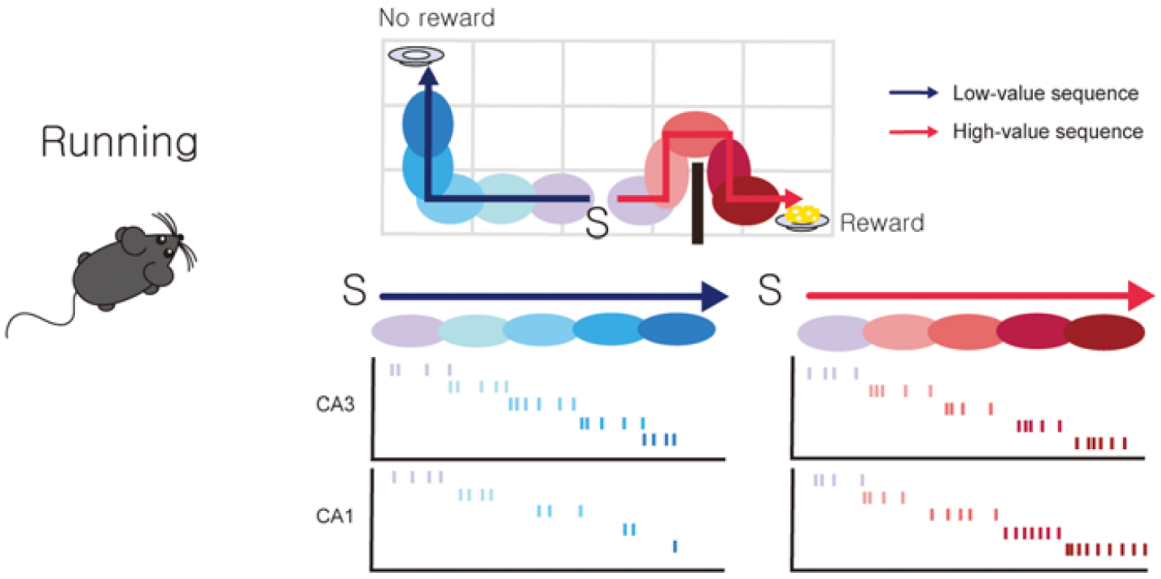

Figure 3.4. CA1 and CA3 activation during low-value and high-value sequences. After several trials of experiencing a path that leads to no reward (low-value sequence) and a path that leads to reward (high-value sequence), with S as the starting point for both paths, neurons in CA1 will increase their firing rate as the mouse gets closer to the reward, and decrease firing rate as it moves away from the reward. CA3 firing rates for both paths are similar. Essentially, it is theorized that CA1 place field activity encodes proximity to reward. Figure is reproduced from M. W. Jung et al. (2018).

Place cell sequence formation at CA3 appears to receive no teaching signal, nor reinforcement signal. In relation to ANN learning rules, this suggests that CA3 specifically employs (auto)-associative unsupervised learning rules for forming sequences of place cell activations. CA1 receives a wider variety of inputs, so it is a bit more difficult to determine what types of learning rules it would use, as it may involve unsupervised learning with just hetero-associations, reinforcement learning if its reward inputs are considered, or possibly even supervised learning if some of its inputs act as a teaching signal, such as an efference copy.

Present effective models of sequence formation in hippocampal function are highly complex. They use realistic spiking conductance-based neuron models to model neuronal network activities of “theta waves” and “sharp wave ripples” in computations requiring massive

computer power and time. We introduce a simple abstracted model that allows for learning sequential presentation of features, as well as backwards replay of those sequences triggered by reward cues. This "Feature Association Matrix" (FAM) model presently uses non-spiking association units, with sensory inputs and a reward input for learning associations, and borrows several other principles from the hippocampus. Presentation of each feature input to an association unit activates a decaying "eligibility trace". During the trace's decay, association between two different inputs can be strengthened or weakened. With presentation of a reward input, a gradient of expected reward is established among linked associations proportional to their relational and temporal proximity to the reward input.

Thus, the association matrix enables memorization and retrieval of sequences of presented sensory cues, particularly when reinforced by reward input. The incorporation of the FAM into an artificial agent, such as the ASIMOV forager, can enable the agent to learn its spatial environment. Elaborating the inputs of the FAM, such as including internal states and proprioceptive feedback, can lead to more complex preferences and behavior in the agent, and may even build the basis for aesthetic creation, such as the sequential construction of nests, burrows, and other structures.

METHODS:

FAM Model

The Feature Association Matrix (FAM) is a simple module for memorization and replay of sequences. For a set of possible inputs or features presented to a forager, such as sensory and reward inputs, the association matrix memorizes the sequence of presentation by appropriately changing associative strength between pairs of inputs. Further, the association matrix assigns

expected reward values to each association, allowing for replay of memorized sequences. Notably, the association matrix is plastic: memorized sequences may be unlearned if they are no longer presented or rewarded, and if there are enough other presentations of competing sequences.

In the FAM, there are three variables for each pair-wise combination of inputs: *Strength, Order,* and *Reward*. Respectively, these determine 1) the degree of correlation (overlap) between the two inputs, 2) the order in which they are received and the relative time interval between them, and 3) the expected reward that their combination provides. As the agent explores its environment (**Figure 3.5**) it receives sensory inputs, and using the FAM it calculates the Strength and Order for each pairwise combination of these inputs, effectively memorizing sequences of associations. Memorized sequences are quite plastic and can easily change unless the agent receives a reward input, perhaps from consuming a rewarding prey item. Once the agent receives reward input, it stabilizes the particular memorized sequences leading to the reward, such that the Strength and Order of associations in this sequence are now less easily changed. Essentially, the Strengths and Orders of associations in the FAM determine a memorized sequence, and the rewards of these associations determine the plasticity and salience of the memorized sequence.

The eligibility trace of an input (**Figure 3.5A**), for instance a sensory or reward input, is a temporary record of the occurrence of that input, which can last longer than the input itself. Therefore, in consuming a rewarding item, the reward input itself may be instantaneous but the eligibility trace of the reward input decays much more slowly, thereby facilitating reinforcement learning for inputs separated by short stretches of time.

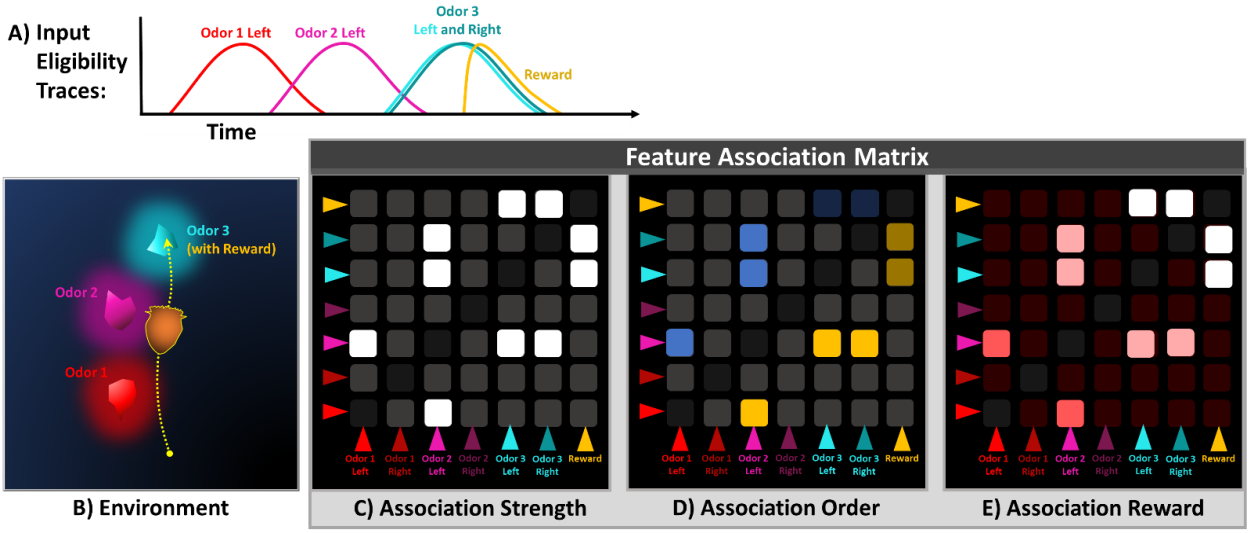

Figure 3.5. The Feature Association Matrix. In this simple case, an agent encounters a few specific inputs, as odors and a reward, in a given environment, and using the feature association matrix, it memorizes the sequence by assigning appropriate strength, order, and reward to the corresponding associations. At the bottom right, the feature association matrix is shown as three different grids, where each arrow in the grids represents an input, with the reward input yellow, and all other arrows as sensory inputs (Odor 1 Left, Odor 1 Right, …, Odor 3 Left, Odor 3 Right). Each box in the grids represents the association between a pair of inputs. **A)** Eligibility traces for the sequence of inputs encountered by the agent. **B)** The given environment that the agent explores following the yellow dotted path, encountering the specific inputs as odors and a reward along the way. **C)** Matrix of association strengths. The strength of an association between two inputs is determined by how correlated the two inputs are in their occurrence. Here, the brightest boxes indicate the strongest associations, indicating that the presented sequence of inputs has been memorized. **D)** Matrix of association Orders. The Order of an association between two inputs is determined by which input occurs first, and the temporal gap between them. The color of each box of the in the grid indicates the relative Order for that association, with grey indicating a zero Order, blue indicating negative Order, and yellow indicating positive Order. Brighter colors indicate Order values of higher magnitude. **E)** Matrix of association Rewards. The Reward of an association between two inputs is determined by how closely the pair of inputs occurs to a Reward input. The color of each box of the association matrix on the bottom right indicates the relative reward value for that association, with the brightest red indicating the highest reward value, and darkest red indicating the lowest reward value. Note that assignment of reward values occurs after the agent has encountered a reward input, which in this case would be the last cyan shape on the yellow dotted path.

The eligibility trace, E_i, of an input i can be expressed as:

$$E_i^{t+1} = \begin{cases} k_E + \dfrac{k_E}{1 + e^{a_E \cdot S_i + b_E}}, & if\ S_i \geq S_{Threshold} \\ d_E \cdot E_i^t, & if\ S_i < S_{Threshold} \end{cases} \quad (3.3)$$

where values for constants k_E, a_E, b_E, d_E, and $S_{Threshold}$ are given in **Table 3.1**, and S_i denotes the intensity of the stimulus input i. When S_i is greater than a sensory threshold for detection, $S_{Threshold}$, the eligibility trace is activated, and when S_i is lower than $S_{Threshold}$, the eligibility trace decays. The *Strength* of an association between input i and input j (**Figure 3.5C**) depends on the overlap of the eligibility traces of input i and input j (**Eqs. (3.4)-(3.6)**):

$$C_{ij}^{t+1} = C_{ij}^t + \Delta C_{ij}, \quad (3.4)$$

$$\Delta C_{ij} = k_C (E_i E_j - E_{Threshold})(E_i + E_j), \quad (3.5)$$

$$Strength_{ij} = \frac{1}{1 + e^{a_S \cdot C_{ij}^t + b_S}}, \quad (3.6)$$

where $Strength_{ij}$ is a logistic function of C_{ij}, and values for the constants k_C, $E_{Threshold}$, a_S, and b_S are provided in **Table 3.1**. If the product of eligibility traces, E_i and E_j, at a certain point in time is less than an established threshold, then Strength is reduced towards the minimum of 0; and if it is greater, then it is increased towards a maximum of 1. Also, changing the association's Strength requires a non-zero sum of the eligibility traces, meaning that at least one eligibility trace (for input i or input j) must be active in order for the Strength to change.

The *Order* of an association between inputs i and j (**Figure 3.5D**) indicates the temporal sequence in which the inputs arrive relative to each other (Eqs. 2.1 and 2.2). For instance, if input j follows input i, Order will increase towards a positive value, and if input j precedes input i, Orer will decrease towards a negative value. If inputs i and j occur at the same time, Order will

be close to zero. Order is simply a scaling of the variable M (**Eq. (3.9)**), which is calculated based on the difference between the eligibility traces of input j and input i (**Eqs. (3.7)-(3.8)**):

$$M_{ij}^{t+1} = M_{ij}^{t} + \Delta M_{ij}\,, \tag{3.7}$$

$$\Delta M_{ij} = \frac{k_M (E_i + E_j)((E_j - E_i) - M_{ij}^{t})}{1 + e^{a_M \cdot Reward_{ij} + b_M}}\,, \tag{3.8}$$

$$Order_{ij} = k_O \cdot M_{ij}^{t}\,, \tag{3.9}$$

where k_M, a_M, b_M, and k_O are constants (**Table 3.1**). To ensure that order and M change only when at least one of the eligibility traces is active, its calculation also depends on the sum of the two eligibility traces. Further, to enhance memorization of a rewarded sequence, an association's M and Order values can be stabilized, when the association has a Reward value assigned to it.

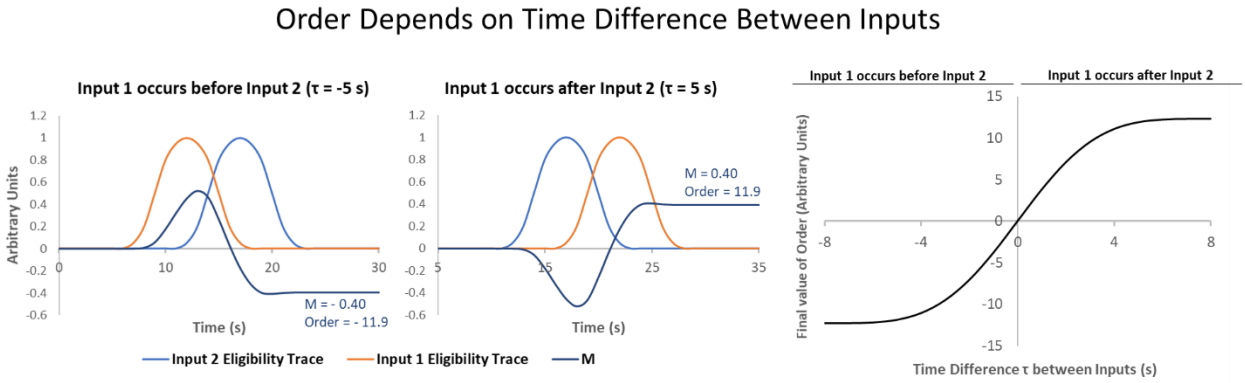

Figure 3.6. The Order between the eligibility traces of two inputs gives an estimate of which input comes first, as well as their timing difference. Order is a scaling of the variable, M, and therefore depends on both the sum and difference of the eligibility traces, and reward value, which is set to 0 for these cases. The first two graphs on the left show the eligibility traces of Inputs 1 and 2, which are delayed relative to each other by a time difference of τ. When Input 1 precedes Input 2, with τ = -5 s, following the end of the eligibility traces, M and Order stabilize at negative values. When Input 1 follows Input 2, with τ = 5 s, these final values of M and Order are positive. The plot on the right shows the final values of Order, for timing difference (τ) values ranging from -8 to 8 s. With sufficient overlap between the eligibility traces, in particular for τ = -3 to 3 s, Order provides a more accurate estimate of the timing difference.

This ensures that a memorized sequence which has provided rewards before is less easily forgotten than memorized sequences that have not provided reward. As shown in

Figure 3.6, M and Order can estimate the time difference between input eligibility traces.

When a reward input is presented, a *Reward* value is assigned to the associations of a memorized sequence, such that the association that occurs closest in sequence to the reward is assigned the highest reward value. This value represents the expected reward of an association. Thus, a gradient of Reward is established for the associations of a memorized sequence, indicating the "proximity" to the reward input (**Figure 3.5E**). Reward value is assigned using the previously calculated Order and Strength of an association. If a reward input is part of the association, then the Reward value is assigned as $Reward_{ij} = k_R \cdot Strength_{ij}$, where k_R is a constant (**Table 3.1**). For each association where neither input directly provides reward, the Reward value is calculated as follows:

$$Reward_{ij} = \begin{cases} Strength_{ij} \cdot \sum_{l=1}^{N} Reward_{lj}, & if \quad Oder_{ij} \geq 0 \\ Strength_{ij} \cdot \sum_{l=1}^{N} Reward_{il}, & if \quad Order_{ij} \leq 0 \end{cases}, \tag{3.10}$$

where N is the total number of different inputs in the FAM. The association's own strength is multiplied by the summed Rewards of all associations that occur right after, as indicated by orders (**Eq. (3.10)**). So if, in a memorized sequence, association #2, with a Reward value of R_2, is the only association that occurs right after association #1, as indicated by the orders, then association #1 will have an assigned Reward value of its own strength times R_2.

ASIMOV Simulation Platform

FAM was incorporated into the agent-based model, **ASIMOV**, in place of the Rescorla-Wagner learning algorithm. Inputs to ASIMOV's forager and the FAM include three different odor inputs, each of which can activate a left or right sensor, and a reward input. ASIMOV's

Incentive variable, which represents the incentive potential of a stimulus, integrates sensory information with innate and learned valences. We express Incentive using the learned Order and Reward values of FAM, and stimulus input intensities as follows:

$$Incentive = \frac{k_I}{N^2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} \Big(Reward_{ij} \cdot (S_i + S_j) \big(T_I - Order_{ij} \cdot S_i \big) \big(T_I + Order_{ij} \cdot S_j \big) \Big), \qquad (3.11)$$

where k_I and T_I are constants (**Table 3.1**), and N is the total number of different inputs in the FAM. After the ASIMOV's agent learns a sequence and Reward values, Incentive values allow it to traverse a learned sequence. For instance, if the agent encounters the first stimulus in a learned sequence, Incentive is initially high and promotes approach.

Table 3.1. FAM Model Constants

Symbol	Description	Value
k_E	Constant for E_i	0.5
$S_{Threshold}$	Threshold for S_i	1e-7
a_E	Exponential constant for E_i	-3.0
b_E	Exponential constant for E_i	7.0
d_E	Decay constant for E_i	0.98
$E_{Threshold}$	Constant for C_{ij}	0.4
k_C	Constant for scaling changes in C_{ij}	4.0
a_S	Exponential constant for $Strength_{ij}$	-0.2
b_S	Exponential constant for $Strength_{ij}$	6.0
k_M	Constant for scaling changes in M_{ij}	0.2
k_O	Constant for scaling $Order_{ij}$	2.0
a_M	Exponential constant for changes in M_{ij}	3.0
b_M	Exponential constant for changes in M_{ij}	-6.0
k_R	Constant for scaling $Reward_{ij}$	10
k_I	Constant for scaling Incentive	3.0
T_I	Constant used for Incentive	1.0

As the agent approaches an overlap of the first and second stimuli of the sequence, Incentive becomes negative, causing the agent to move away from the first encountered stimulus and towards the next one.

For all simulations, we fixed ASIMOV's Reward Experience and Satiation variables at values 0 and 0.10, respectively, in order to see how the FAM's learning algorithms affect the agent's approach/avoidance and foraging behavior. Both ASIMOV and the FAM module are implemented in the graphic, agent-based programming language, NetLogo (Wilensky, 1999).

RESULTS:

To explore how the FAM affects the ASIMOV forager's foraging and navigation behavior, we set up a virtual environment with three different overlapping odor sources, as pictured in **Figure 3.7**. During the simulation, ASIMOV's forager initially encountered the sequence of odor sources, Odor 1, Odor 2, Odor 3, receiving reward input only when it encountered the source of Odor 3 at around 200 ticks. Strengths and Orders of the FAM were established as the forager encountered the odor sources from t = 0 to 200 ticks. FAM Reward values were established only when the forager received reward input at ~200 ticks. At later times, the FAM showed almost no change in Strength, Order, or Reward values. Notably, after first encountering the odor sequence, the forager no longer had direct contact with Odor 1 and Odor 2 sources, and exclusively went for the Odor 3 source. This is because only the Odor 3 source provided reward throughout the simulation. As seen after t = 200 ticks (**Figure 3.7**), the forager often stayed in close proximity to Odor 1 and Odor 2 sources, navigating along the edges of the odors in either direction, until it reached Odor 3 source. This simulation suggests that with FAM, the forager is able to learn a sequence of stimuli that it encounters in the

environment, and use the learned associations to navigate the environment in ways that provide more reward.

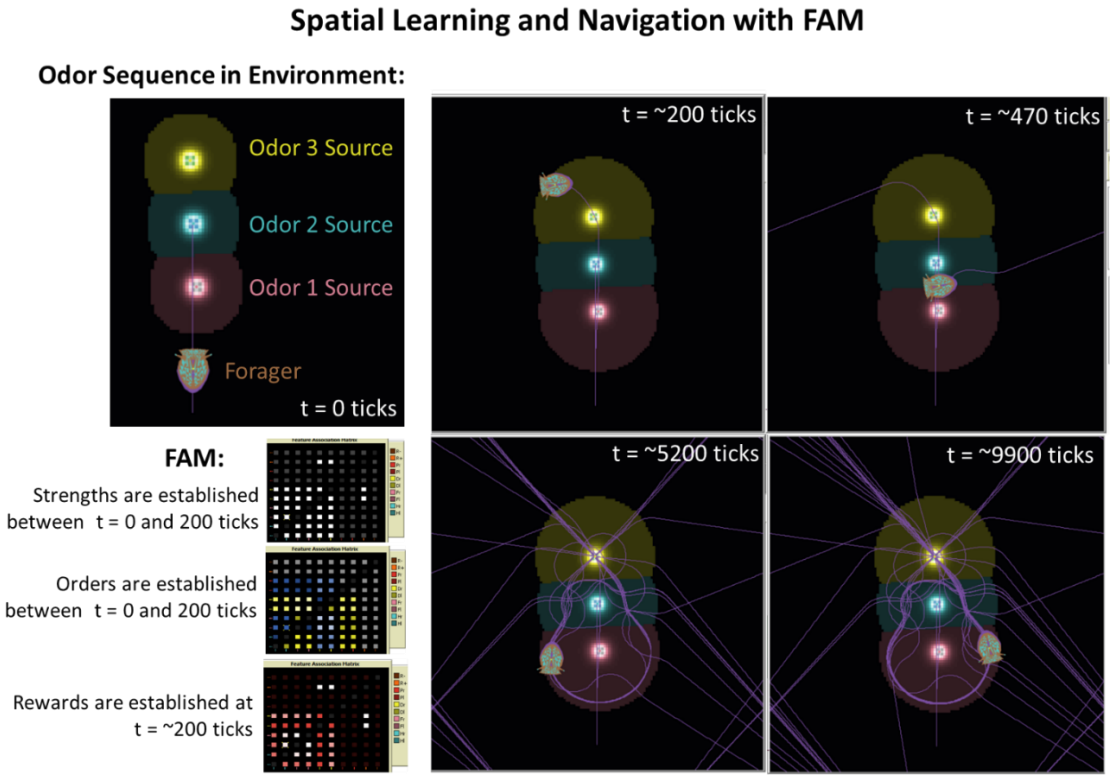

Figure 3.7. ASIMOV's forager with FAM, showing spatial learning and navigation. In a virtual environment, the forager encounters a sequence of odor sources, Odor 1, Odor 2, and Odor 3 for t = 0 to 200 ticks. The purple lines indicate the forager's previous movement. At t = ~200 ticks, the forager encounters Odor 3 Source and receives a reward input, establishing Reward values in its FAM. After 200 ticks, the forager exclusively comes in direct contact with the Odor 3 Source, as it is the only one that provided reward throughout the simulation.

The FAM may also enable the forager to use second-order conditioning mechanisms (Hawkins et al., 1998). Second-order conditioning is a type of associative learning where a neutral stimulus is paired with a conditioned stimulus to elicit a conditioned response. Notably, second-order conditioning appears to be impaired with hippocampal lesions (Gilboa et al., 2014). Using Presentation Mode described previously in ASIMOV Supplemental Methods, in which the forager is immobilized but retains free turning responses, we presented a temporal sequence of odors: Odor 1, Odor 2, and Odor 3, followed by a reward input (**Figure 3.8**).

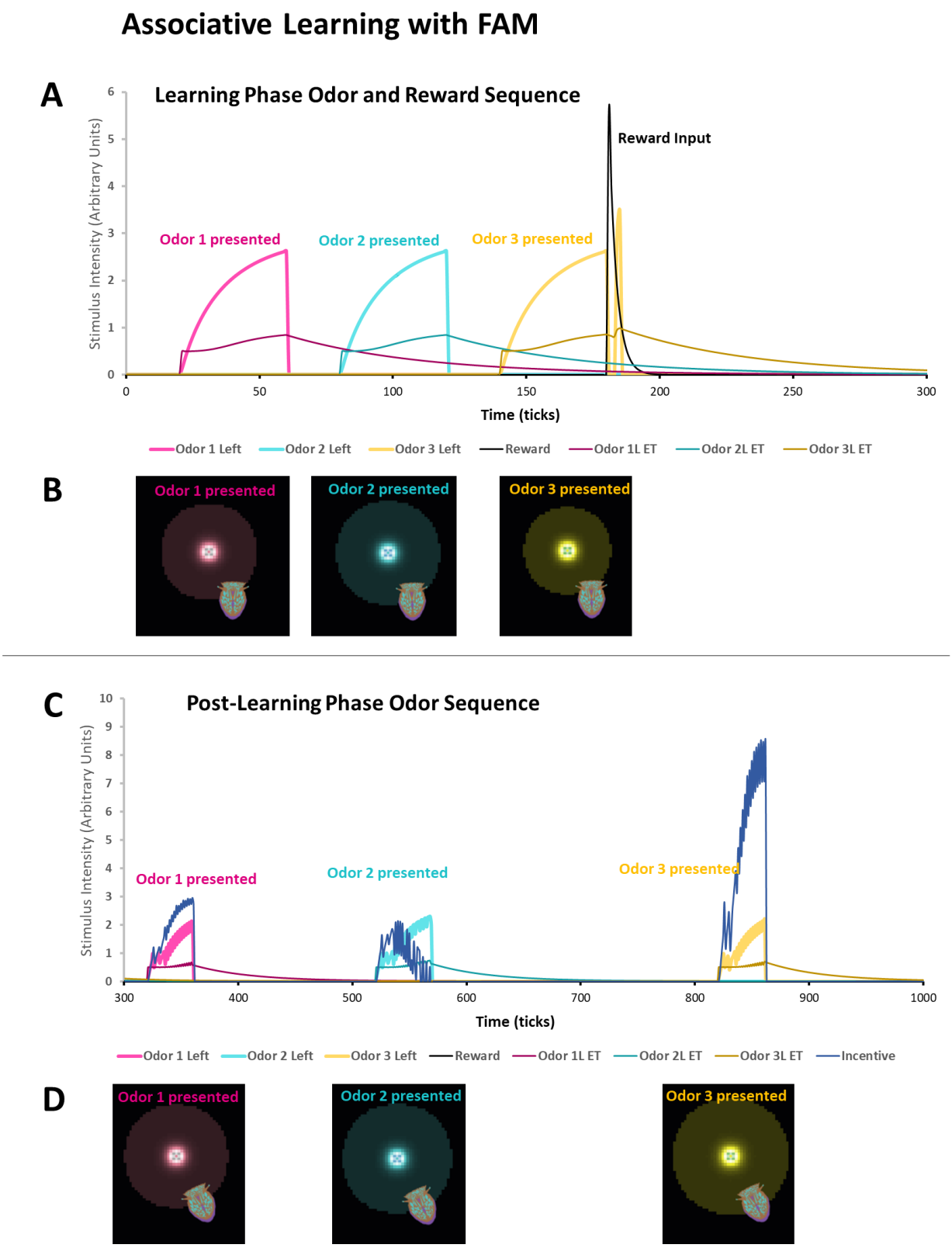
Associative Learning with FAM
A
Learning Phase Odor and Reward Sequence
Stimulus Intensity (Arbitrary Units)
Reward Input
Odor 1 presented
Odor 2 presented
Odor 3 presented
Time (ticks)
Odor 1 Left
Odor 2 Left
Odor 3 Left
Reward
Odor 1L ET
Odor 2L ET
Odor 3L ET
B
Odor 1 presented
Odor 2 presented
Odor 3 presented
C
Post-Learning Phase Odor Sequence
Stimulus Intensity (Arbitrary Units)
Odor 1 presented
Odor 2 presented
Odor 3 presented
Time (ticks)
Odor 1 Left
Odor 2 Left
Odor 3 Left
Reward
Odor 1L ET
Odor 2L ET
Odor 3L ET
Incentive
D
Odor 1 presented
Odor 2 presented
Odor 3 presented

Figure 3.8

Figure 3.8. FAM allows for higher-order associative learning in ASIMOV's forager. Using Presentation Mode, where the forager is immobilized but can still freely turn, a temporal sequence of odors is presented: Odor 1, Odor 2, Odor 3, followed by a reward. A) During the learning phase, there is no temporal overlap in presentation of odors, but there is overlap in eligibility traces (ETs). B) shows that the forager does not turn towards the presented odors in the initial learning phase. C) During a second set of the same odor presentations, where even the eligibility traces do not overlap, Incentive is increased during each odor presentation. Notably, (D) shows that the forager turns towards the presented odor sources, showing that it has learned the associations in the first presented set of inputs.

The forager does not turn towards the odor sources during these initial presentations (**Figure 3.8**B), and there is no temporal overlap in odor presentations. However, there is overlap in the eligibility traces of these odors (**Figure 3.8A**). During the first phase of odor presentations, the forager establishes Strengths and Orders of associations using the eligibility traces, and establishes Reward values during the reward input. In the second phase of the same odor presentations (**Figure 3.8C**), there is no overlap in eligibility traces. Notably, the Incentive spikes during the odor presentations because of the previously learned associations. This causes the forager to make an appetitive turn towards the odor source in each case (**Figure 3.8D**).

While **Figure 3.8** does not explicitly demonstrate the typical second-order conditioning paradigm, another simulation shows similar increases in Incentive and appetitive turning behavior (**Figure 3.9A-B**), with stimuli presented in closely occurring pairs during the learning phase: Odor 3 then reward input, Odor 1 then Odor 3, and Odor 2 then Odor 1. In this case, even though the reward is given early on, after the forager's first encounter with Odor 1, it is still able to assign expected rewards to subsequently formed associations (**Figure 3.9C**), due to Strength and Order calculations. Interestingly, Incentive values decrease with repeated presentations of an odor (**Figure 3.9A**). This is because after the first 200 ticks, these odors are repeatedly presented without reward, thus decreasing expected reward values as seen in the much fainter FAM

expected reward values for t = 1254 ticks (**Figure 3.9C**), which affects Incentive calculations. This allows the forager to ignore stimuli that are no longer rewarded.

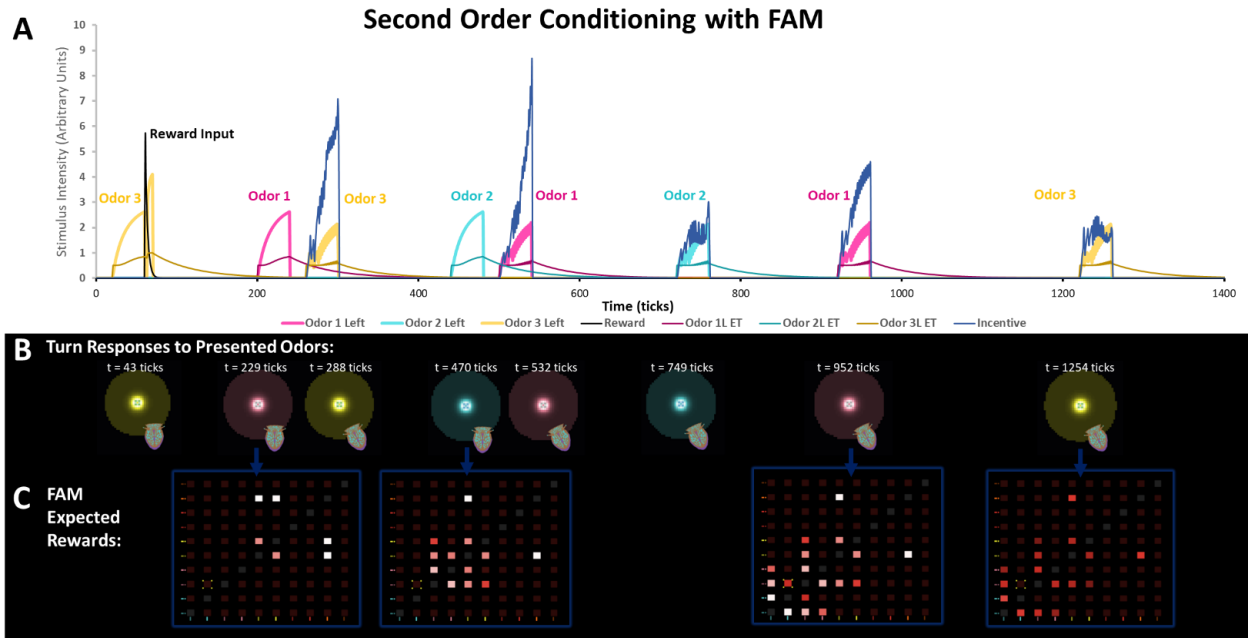

Figure 3.9. FAM enables second-order conditioning for ASIMOV's forager. **A)** Similar to **Figure 3.8** simulation, the forager is able to form associations and increase stimulus-specific Incentive values as it encounters a temporal sequence of odors. In this case the stimuli are presented as closely-occurring pairs during the first 600 ticks: Odor 3 then reward input, Odor 1 then Odor 3, and Odor 2 then Odor 1. Note that during this period, there is no temporal overlap in presentation of odors, but there is overlap in eligibility traces (ETs). **B)** Turning responses corresponding to odor presentations in (A). After encountering all pairs of inputs, the forager shows appetitive turns towards all the odors. **C)** Visualization of FAM expected rewards at specific time points as indicated by the blue arrows and corresponding turning responses in (B).

DISCUSSION:

The Feature Association Matrix (FAM) is a simple model for memorization and replay of sequences. In particular, it establishes a memorized sequence by appropriately changing associative variables, Strength and Order, for each pair-wise combination of inputs. Sequence salience is established during reward input, using the previously learned associations to assign expected Reward values for each input pair. We show that this simple model, when incorporated

into the ASIMOV agent, enables spatial learning and navigation (**Figure 3.7**), as well as higher-order associative learning, such as second-order conditioning (**Figure 3.8**-**Figure 3.9**).

FAM is an abstraction of physiological circuits that learn sequences of sensory cues, such as those responsible for place field coding in the hippocampus (**Figure 3.3**). Each association of the matrix is analogous to the synapses of place cells in region CA3 of the hippocampus, which make up a recurrent auto-associative network responsible for the learning and formation of navigation sequences (Hasselmo et al., 1995; Van Strien et al., 2009). The assignment of reward value to sequences formed in CA3 may occur through CA1 of the hippocampus (M. W. Jung et al., 2018), which receives dopaminergic input from midbrain structures, such as substantia nigra compacta and ventral tegmental area. Notably, the processing of reward, as encoding of reward magnitude or reward prediction error, is uniquely associated with the reverse replay of sequences, rather than forward replay (Ambrose et al., 2016). Similarly, the FAM assigns reward values through a reverse traversal of a learned sequence, using previously established strength and order values. FAM can thus be characterized as an artificial neural network-based memory model with unsupervised and reward-based learning rules.

For a simple example of a neuronal circuit analogue of the Feature Association Matrix, two sensory inputs (Odor 1, Odor 2) and one reward input (Reward 1) are used in a circuit that describes association matrix function (**Figure 3.10**):

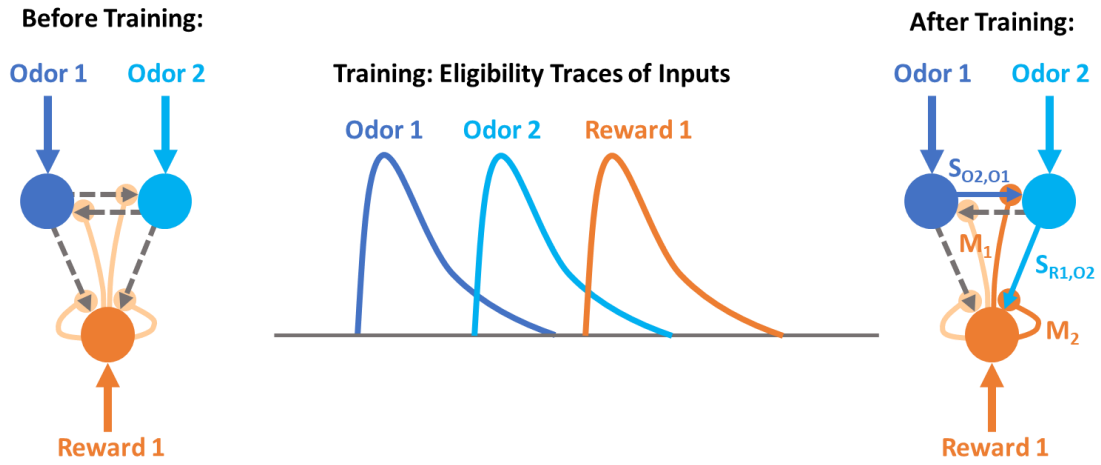

Figure 3.10. A neuronal circuit analogue for the association matrix. In this example there are two sensory inputs, Odor 1 and Odor 2, and one reward input, Reward 1. Before training, no associations are formed yet (gray connections), and no reward values have been assigned yet (light orange connections). During training, the sequence of three inputs is presented as Odor 1, Odor 2, Reward 1. The sequential overlap of the eligibility of these three inputs increases the strengths of the associations $S_{O2,O1}$ and $S_{R1,O2}$ (dark orange and solid blue connections) after training. Following reward input , there is also an assignment of reward value to the associations $S_{O2,O1}$ and $S_{R1,O2}$ through the strengthening of heterosynaptic connections M_1 and M_2.

Before training, all synapses, or associations are weak. Let O_1, O_2, and R_1, denote the neurons receiving input Odor 1, Odor 2, and Reward 1, respectively. During training, inputs are presented in the sequence Odor 1, Odor 2, Reward 1, such that their eligibility traces overlap. The synapse from neuron O_1 to O_2 is strengthened, followed by strengthening of the synapse from O_2 to R_1 neuron. The plasticity mechanism by which synapses are strengthened is analogous to spike timing dependent plasticity (STDP); this is expressed in the code through the association's "Strength" and "Order" variables. Thus, after training, the colored synapses $S_{O2,O1}$ and $S_{R1,O2}$ represent the associations that are formed. The neurons that receive reward (like R_1) heterosynaptically facilitate these synapses. In this case, the modulatory synapses of neuron R_1 are strengthened through positive feedback. If the reward neuron reactivates synapses that have already been strengthened ($S_{O2,O1}$ and $S_{R1,O2}$), then this will essentially replay the memorized

sequence. This replay can then strengthen the reward neuron's synapses, M_1 and M_2, on the basis of "shortest path", such that M_2 is strengthened more than M_1. This would establish a reversible sequence via the reward neuron, and this is reflected in the code in the assignment of a reward value to each association.

Similarities to Other Models

One of the earliest examples of an auto-associative network for sequence memorization is an auto-associative correlation matrix memory (CMM) (Cutsuridis & Wennekers, 2009; Kohonen, 1972). The CMM is based on a correlation matrix of component input signals, and uses Hebbian modification of connections to memorize patterns among binary input elements. This could be considered one of the simplest ANNs that uses a CA3-like architecture as well as Hebbian learning. However, the CMM does not learn sequences of inputs, and uses very simple binary inputs. Unlike CMM models, the FA Matrix does not rely on binary stimulus events, and also takes into account the order and expected reward of each pair-wise combination of input elements.

One of the more advanced models that mimics hippocampal structure is that of Lawrence et al. (Lawrence et al., 2006), which uses both auto-associative and hetero-associative connections between two modules, A and B, pictured below (**Figure 3.11**):

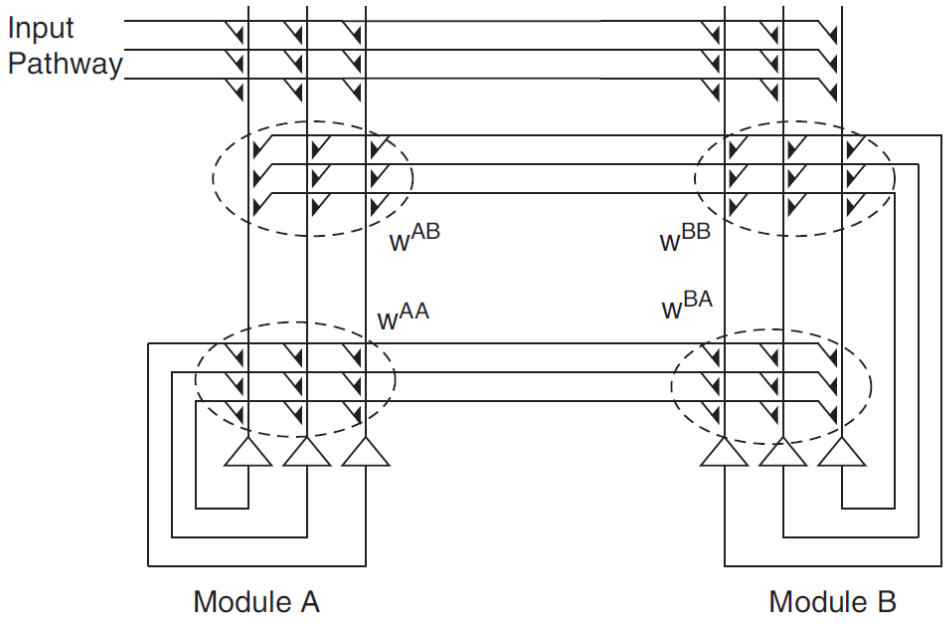

Figure 3.11. Network architecture of a hippocampal-like model for sequence learning developed by Lawrence et al. (2006) with strong auto-associative weights and slow hetero-associative weights. Figure reproduced from Lawrence et al. (2006).

This model utilizes CA1 and CA3-like mechanisms and architecture: it uses recurrent connections, spiking LIF neurons, and Hebbian learning with a delay component, to form robust sequence memory. The auto-associative weights are used for pattern completion, while the hetero-associative weights drive the state of the system from one pattern to the next. This model shows improved performance over the Hopfield model, another widely-used auto- and hetero-associative sequence-learning, by using slower, or delayed, hetero-associative synapses which allows the network to retain more of a short-term memory effect. The FAM has a different and simpler architecture, but uses more complex learning rules to establish associations. In contrast to the model of Lawrence et al., and most other neural network-based spatial memory models, the basic units used in the FAM are primarily analogs of synapses, rather than neurons (Madl et al., 2015). Similar to the models of Barrera et. al (2011) and Strösslin et al. (2005), the FAM uses both Hebbian-like and reinforcement learning to assign expected rewards, but in addition it

employs a simple timing dependent plasticity, similar to spike-timing-dependent plasticity (STDP), to determine the order in which inputs occur in a sequence. In contrast to spiking models of hippocampal memory (Buzsáki, 1989; Cutsuridis & Wennekers, 2009; Kunec et al., 2005), the FAM does not have any spiking units and does not use any conductance-based neuron models. While this means that the FAM is less biophysically realistic, it also means that the computations for FAM are simpler, take less time and computing memory, and may be easier to expand to a much greater scale than most biophysical memory models.

Learning Precise Timing through NMDAR Conductance Mismatch in a Computational Model of Developmental Synaptic Plasticity

It has been repeatedly shown that neuronal synapses are able to change their synaptic strengths by regulating receptors in an activity-dependent manner, as seen in learning and memory formation. Likewise, the idea of altering synaptic weights to learn and achieve specific output patterns has been heavily incorporated in artificial neural networks (ANNs) for decades. However, a largely unexplored question in the field of neurobiology is whether neuronal synapses are capable of learning by altering the *timing* with which they receive synaptic inputs. For instance, many models of learning, neural development, and plasticity focus on changes in synaptic weights; however, few explore changes in the time delay of input signals. Barn owls and electric fish show extreme temporal precision of sensory delay lines that may be explained by genetic factors, axonal guidance cues, and synaptic pruning. *However, there is a high probability that the timing of these delay lines is fine-tuned in development to allow a neuron to learn and adjust the relative timing with which it receives input signals.* The existence of such a mechanism would not only increase the computational powers and complexity of biological synapses, but it would likewise introduce new avenues for developing ANNs that are able to learn sequences of inputs at synaptic and single neuron levels, instead of using entire networks for sequence memorization.

This paper introduces such a hypothesis. In a computational single-compartment model, SITDL (Synaptic Input Time Difference Learning), the N-methyl-d-aspartate receptor (NMDAR) expression in a silent synapse is affected by the difference in timing between two signals that the synapse receives: a synaptic glutamate signal and a dendritic voltage signal that

travels past the synapse. The difference induces realistic NMDAR modifications in allosterism and/or composition. Specifically, NMDARs have gates that are independently activated by glutamate binding and depolarization. The SITDL model stipulates that timing of the NMDAR glutamate gate activation changes, based on the timing difference between voltage and glutamate signals. This enables the silent synapse to achieve greater coincidence between glutamate and voltage gate activations, and effectively encode the signals' timing difference in the NMDAR glutamate gate activation time. This proposed mechanism can account for how delay line tuning is possible to such extreme precision. Moreover, it questions the nature of memory formation in hippocampal synapses, whose activity depends heavily on NMDAR plasticity. While the biological plausibility of such a mechanism is yet to be tested, SITDL would be a significant mechanism for spiking ANNs in machine learning, which up until now, have only employed changes in synaptic weight rather than synaptic timing.

INTRODUCTION:

Synaptic plasticity is a process in which synaptic properties, such as the efficacy of synaptic transmission, are modified, often through activity-dependent changes in expression of post-synaptic receptors or in pre-synaptic neurotransmitter release. Most studies of synaptic plasticity, in both physiological systems and ANNs, focus on the strengthening or weakening of synapses, while changes in synaptic timing are often left unexplored. In physiological networks, there are many different forms of synaptic plasticity, including the widely expressed NMDAR-dependent long-term potentiation (LTP) and long-term depression (LTD), which may play an important role in learning and memory processes (Cooke & Bliss, 2006).

NMDAR-dependent LTP and LTD

In NMDAR-dependent plasticity, strongly correlated and coincident pre-synaptic and post-synaptic activity leads to synaptic strengthening, or LTP, and weakly correlated pre- and post-synaptic activity leads to synaptic weakening, or LTD. Both cases involve regulation of post-synaptic α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic receptors (AMPARs) through NMDAR activation (Cooke & Bliss, 2006). While both AMPARs and NMDARs are heteromeric cation channels that bind glutamate, NMDARs are particularly unique, since for activation they require binding of glycine and glutamate, which alter the conformation of the NMDAR channel, as well as local depolarization, which expels the magnesium blockade (Hassel & Dingledine, 2012; Lester et al., 1990; Lüscher & Malenka, 2012). Essentially, the NMDAR acts as a coincidence detector, as it has both a glutamate gate and voltage gate that require near-coincident activation for the NMDAR channel to fully open. Note that throughout this paper, mentions of glutamate binding and activation of NMDAR's glutamate gate will imply the coincident binding of the co-agonist glycine. Once NMDARs are activated, they permit cation flow, significantly consisting of Ca^{2+} influx, which causes a slow depolarization. If this Ca^{2+} influx is significantly large, often resulting from coincident pre- and post-synaptic depolarizations, then through a series of chemical cascades involving CaM kinase II, AMPARs can be phosphorylated and inserted into the post-synaptic membrane, resulting in LTP and synaptic strength increase. If Ca^{2+} influx is small, then this can lead to dephosphorylation of AMPARs and their removal from the post-synaptic membrane, causing LTD and decreasing the synaptic strength.

Notably, the numbers of AMPARs and NMDARs in synapses are estimated to range from tens to a few hundred (Kennedy et al., 2005). While these numbers depend on multiple factors, including brain region, developmental stage, and synaptic size, it seems that AMPARs vary more widely in number, typically from zero to fifty receptors per synapse (Hassel & Dingledine, 2012), than NMDARs, which are present even at "silent" synapses that do not generate any detectable excitatory post-synaptic potentials (EPSPs) in response to neurotransmitter release.

NMDAR-dependent LTP is strongly implicated in learning and memory, including spatial working memory, stability of spatial representations, and associations established by reward learning (Bannerman et al., 2008; Bannerman et al., 2014; Bethus et al., 2010; Kentros et al., 1998). In the hippocampus, NMDAR-dependent plasticity is particularly important for spatial representations, as genetic knockdown of NMDARs in principal CA1 neurons has been shown to disrupt hippocampal oscillations and selectively impair sequence-based place cell representations (Cabral et al., 2014).

NMDAR Dynamics and Expression

Many studies emphasize NMDARs' role in changing synaptic strengths and their importance in memory, however, few have examined NMDARs' possible role in changing synaptic timing. I propose that NMDARs are indeed important for changing synaptic timing in an activity-dependent manner, as NMDAR dynamics can be regulated in many different ways, and NMDAR composition can change over time, with different subunits conferring distinct activation rates to NMDARs.

The dynamics of an NMDAR, such as glutamate binding and channel opening rates, depend on the subunits that compose it. Each NMDAR is made up of two obligatory GluN1 subunits, and two additional subunits of GluN2 or GluN3. GluN2 and GluN3 subunits have distinct subtypes, which can determine NMDAR gating properties (Paoletti et al., 2013). We are interested in the GluN2 subtypes, GluN2A and GluN2B, as they are the predominant subunit types found in hippocampus and in many brain structures that rely on precise timing, and they have very different expression profiles in development. GluN2A-containing NMDARs typically have faster activation, faster deactivation, and higher affinity for glutamate than GluN2B-containing NMDARs (Paoletti et al., 2013). Notably, computational modeling of NMDAR subtype activation suggests that diheteromeric GluNR1/GluNR2A NMDARs have a rise time to peak activation of about 7 ms, while diheteromeric GluNR1/GluNR2B NMDARs have a significantly longer rise time of approximately 50 ms (Singh et al., 2011). This suggests that changes in NMDAR subunit composition may effectively change how quickly glutamate activates the NMDAR glutamate gate, thus changing the timing of NMDAR-mediated Ca^{2+} influx at a synapse.

There is significant regulation of NMDAR dynamics through ligand binding and modifications (Hassel & Dingledine, 2012). This includes a minimum of six distinct ligand binding sites on the NMDAR that can affect the probability of NMDAR channel opening, such as a polyamine regulatory site, and recognition sites for agonists and different ions. NMDARs also interact with proteins in the post-synaptic density. The anchoring of NMDARs to the post-synaptic density protein, PSD-95, through their GluN2 subunits, stabilizes their expression in the post-synaptic membrane and also couples them to intracellular signaling systems involving calmodulin (Hassel & Dingledine, 2012). The binding of the Ca^{2+}-calmodulin complex to an

NMDAR results in a Ca^{2+}-dependent reduction of the NMDAR's channel opening frequency and channel open time. Furthermore, NMDAR activity regulates casein kinase 2 phosphorylation of GluN2B subunits, which disrupts the subunits' interactions with PSD-95 and decreases the GluN2B surface expression in neurons (H. J. Chung et al., 2004). Interestingly, GluN2B subunits also appear to be more mechanosensitive than GluN2A subunits, such that mechanical stretch induces greater Ca^{2+} influx in NMDARs containing GluN2B, possibly due to a reduction of efficacy of the Mg^{2+} block (Singh et al., 2012). The wide diversity of ways in which NMDAR expression and dynamics can be modulated, through binding, modifications, and changing subunit composition, provides a wealth of opportunity for activity-dependent changes in synaptic current timing through NMDARs.

NMDAR subunit types have different spatiotemporal expression profiles (Paoletti et al., 2013), which result in distinct synaptic dynamics in different brain regions and at different times in development. For instance, GluN2B expression is highest in the mouse brain early in development, and after the first post-natal week, expression levels drop, with most GluN2B expression becoming restricted to the forebrain. GluN2A expression levels, on the other hand, are lower after birth, and grow during the course of development, with most neurons in the adult mouse brain expressing GluN2A. Notably, neonatal synapses can experience rapid activity-dependent bidirectional switching in subunit composition, between GluN2A and GluN2B containing NMDARs, on the order of seconds (Bellone & Nicoll, 2007). The hippocampus, a mammalian structure important for learning and memory, is one of the few known brain regions that has mechanisms for rapid bidirectional GluN2A and GluN2B subunit switching even in adulthood (S.-C. Jung et al., 2008).

Distinct NMDAR expression is implicated not only in brain structures important for memory, but in other systems that rely on precise timing, such as in mammalian Calyx of Held and avian auditory brainstem (Futai et al., 2001; Joshi et al., 2007; Tang & Carr, 2007), important for inter-aural time and intensity difference calculations in sound localization, as well as in weakly electric fish relay cells, which are important for precise temporal regulation of the jamming avoidance response (Bekkers & Stevens, 1990; Zakon et al., 2002). Notably, these delay line systems can resolve temporal disparities in the microsecond range, which may arise at the level of individual neurons. For instance, single neurons of the pre-pacemaker nucleus in a weakly electric fish are sensitive to temporal disparities as small as 1 μs, and their signaling and precision may involve NMDARs (Kawasaki et al., 1988; Spiro et al., 1994). While there are relatively few studies on changes in NMDAR subunit composition during the development of these delay line systems, it has been shown that NMDAR composition does change in avian auditory brainstem development, in particular, with GluN2A subunits replacing GluN2B subunits in chicken cochlear nuclei, after onset of hearing (Tang & Carr, 2007).

SITDL Hypothesis

Much is still unknown about NMDARs, such as the intricacies of their dynamics and regulation, the roles of different subunits, and their specific role in memory formation. Thus, I propose a hypothetical model, Synaptic Input Time Difference Learning (SITDL), of how activity-dependent changes in fast and slow NMDAR expression can affect synaptic current timing. *In particular, I assume that synaptic NMDAR structure depends on differences in voltage and glutamate gate conductances, and that matching of these conductances provides optimal* Ca^{2+} *influx through NMDAR gates.* This process is driven by changes in subunit composition of

the NMDARs, which determine glutamate gate characteristic activation time. We provide a specific example of this hypothesis in **Figure 3.12**.

The left panel of **Figure 3.12** shows an example of a silent synapse that starts with a majority of slow NMDARs, which have a slower activation time in response to glutamate than fast NMDARs. The slow NMDARs can be considered as NMDARs containing GluN2B subunits, and fast NMDARs as containing GluN2A subunits. The synapse receives a synaptic glutamate signal followed by a dendritic voltage signal, whose peaks occur several milliseconds apart. This difference in timing causes a difference in NMDAR voltage and glutamate gate conductance. I propose that as the synapse develops and experiences the same timing between voltage and glutamate signals, its NMDAR population may change due to the difference in glutamate and voltage gate conductances.

In the right panel of **Figure 3.12**, fast NMDARs replace slow NMDARs until there is a minimization of the difference in glutamate and voltage gate conductance, thus aligning the overall NMDAR current peak with the peak of the voltage signal. This effectively allows the synapse to learn the timing difference between the voltage and glutamate signals, encoding it in the NMDAR population's glutamate activation time, which depends on the numbers of fast and slow NMDARs. Then, stabilization of a synaptic NMDAR population, such that there are no more removals and insertions of fast or slow NMDARs, possibly due to PSD-95 anchoring, would mean a stable memory of the timing difference. Note that silent synapses are particularly useful for this SITDL mechanism, as glutamate signals alone do not cause depolarization of the synapse, thus providing a greater degree of independence between NMDAR glutamate gate and voltage gate activation. I explore the hypothesis of SITDL further, and provide a computational single-compartment model for it in the Methods section.

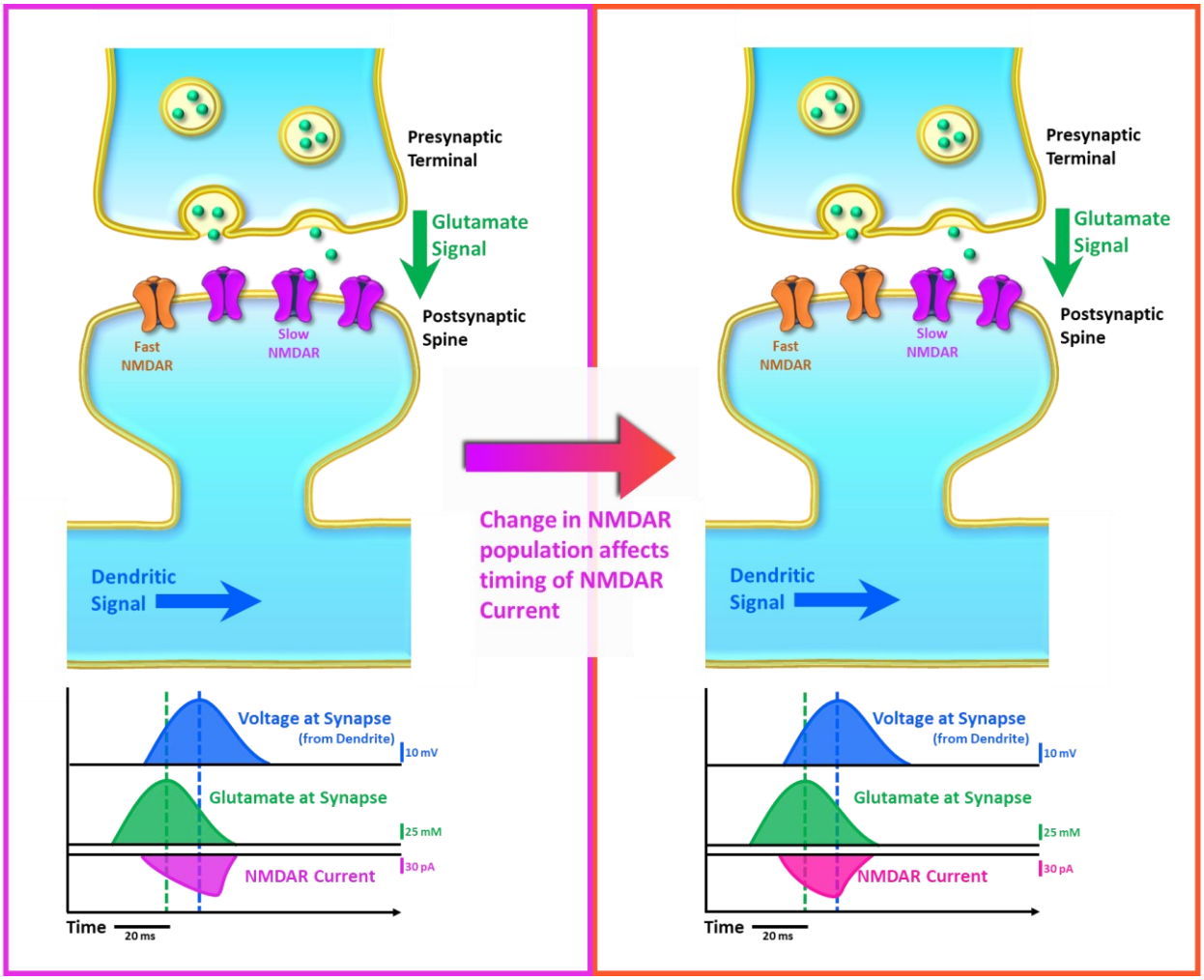

Figure 3.12. An example of the SITDL hypothesis, showing how activity-dependent changes in NMDAR population can affect synaptic current timing. **Left**: A developing silent synapse starts with a majority of slow NMDARs, which have a slower glutamate gate activation than fast NMDARs. The synapse receives a glutamate signal followed later by a dendritic voltage signal, causing a difference in the activation of NMDARs' voltage and glutamate gates. **Right**: We propose that as the synapse develops and experiences the same timing between voltage and glutamate signals, its NMDAR population may change due to the difference in glutamate and voltage gate activation. In this case, fast NMDARs replace slow NMDARs until there is a minimization of the gate activation difference, thus aligning the overall NMDAR current peak with the peak of the voltage signal. This idea and mechanism of SITDL may allow the developing synapse to learn the timing difference between the voltage and glutamate signals by encoding it in the NMDAR population's glutamate gate activation time.

METHODS:

Input Signal Generation

To construct dendritic and glutamate signals for testing the plasticity model, periodic input signals, $S_N = S_N(t)$ and $S_L = S_L(t)$ (**Figure 3.13**), were respectively generated using a gaussian-like shape for the spikes in voltage, and right-skewed shape for synaptic glutamate concentration (Einolghozati et al., 2011; Sargent et al., 2005; Ventriglia & Di Maio, 2000) :

$$S_N(t) = e^{-\left(\frac{(t-t_{N0})}{\sigma_N}\right)^2}, \tag{3.12}$$

$$S_L = (t - t_{L0})e^{-\alpha_L(t-t_{L0})}, \tag{3.13}$$

$$S_N = S_N / S_{NMax}, \tag{3.14}$$

$$S_L = S_L / S_{LMax}, \tag{3.15}$$

$$I_D(t) = I_0 + S_N(t), \tag{3.16}$$

$$S_{Glu}(t) = S_L(t), \tag{3.17}$$

where $S_N = S_N(t)$ is a gaussian-like function that determines the shape of the voltage spike, t_{N0} is picked from a pre-determined repeating sequence of durations, and determines the inter-spike interval, and the constant σ_N determines the width of the spike. $S_L = S_L(t)$ determines the right-skewed shape of the glutamate spike, the constant α_L determines the width and decay of the spike, and t_{L0} is simply shifted relative to t_{N0} , with $t_{L0} = t_{N0} - 1/\alpha_L$, to ensure that the spike peaks of both $S_N(t)$ and $S_L(t)$ occur at the same exact time (**Figure 3.13**). The input signals, S_N and S_L were normalized by their respective maximums, S_{NMax} and S_{LMax} (**Eqs. (3.14)-(3.15)**). The dendritic signal, $I_D = I_D(t)$, depends on I_0, a constant current, and the signal S_N. The glutamate signal, $S_{Glu}(t)$ is equivalent to $S_L(t)$. Values for all constants are given in **Table 3.1**.

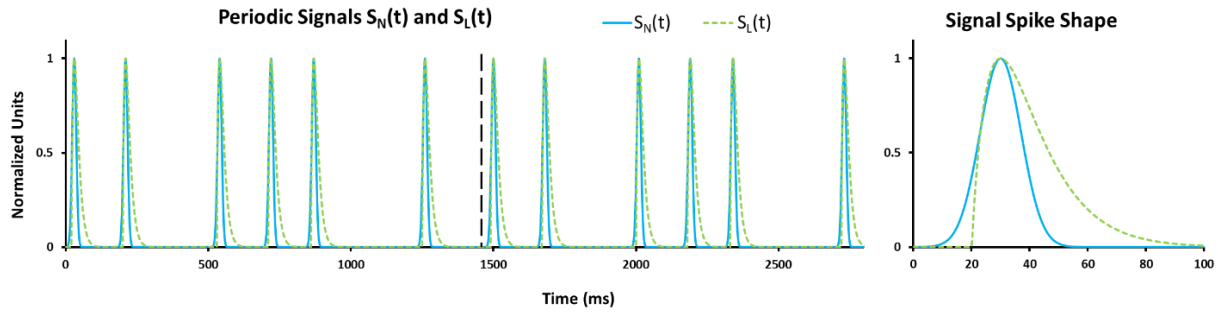

Figure 3.13. Periodic signals $S_N(t)$ and $S_L(t)$ used for voltage and glutamate signal generation, respectively. Left: the timing of peaks is the same for both signals and follows a repeating pattern, as indicated by the black dashed line. Right: $S_N(t)$ spikes have a gaussian-like shape, while $S_L(t)$ spikes have a right-skewed shape.

NMDAR Voltage and Glutamate Gate Dynamics of a Single Synapse

NMDARs are important in learning, memory, development, as well as regulation of synaptic properties at the post-synaptic membrane. Each NMDAR is a heteromeric cation channel with a glutamate-activated gate and voltage-activated gate. Activations of the glutamate gate or voltage gate are largely independent, and alone either does not open the NMDAR channel completely. However, it will cause a change in conformation or channel structure. With near-coincident activation of both gates, the NMDAR channel opens, permitting the influx of Ca^{2+}.

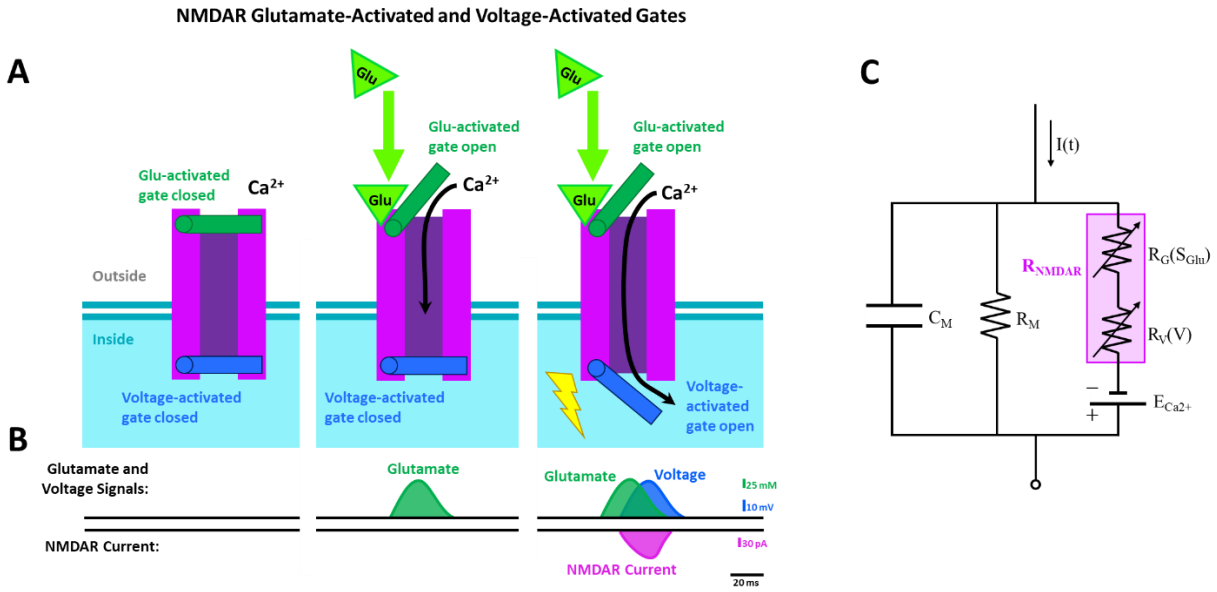

Figure 3.14. NMDAR glutamate and voltage gates. Each NMDAR has a glutamate-activated gate, that is opened upon binding glutamate, as well as a voltage-activated gate, in the form of a Mg^{2+} blockade, which opens when there is sufficient local depolarization of the NMDAR. **A)** The leftmost image shows that at resting membrane potential, and when no glutamate is bound, the NMDAR channel is closed. In the middle channel, glutamate binds to an NMDAR that is still at resting membrane potential, causing the glutamate gate to open, and the voltage gate to remain closed. This means that the NMDAR channel is not open, however, there is still a change in conformation that may allow Ca^{2+} to partially enter the channel, but not be able to enter into the cell. The next image shows near-coincident glutamate binding and depolarizations, which cause opening of the NMDAR's glutamate gate and voltage gate, respectively. This results in the NMDAR channel being fully open, which permits Ca^{2+} to enter the cell. **B)** Graphs of synaptic glutamate concentration, voltage from dendritic signals, and NMDAR currents corresponding to the three images above in A. The leftmost graph shows no glutamate or voltage signal, and the middle graph shows only a glutamate signal, so there is no activation of the NMDAR current in either case. The next graph shows a glutamate signal followed by a dendritic voltage signal, whose coincidence activates the NMDAR current. **C)** Circuit diagram of a post-synaptic spine containing only NMDARs. Specifically, we can think of an NMDAR (R_{NMDAR}) as a resistor that is an in-series composition of its glutamate gate (R_G) and voltage gate (R_V), which are shown as variable resistors that depend on the glutamate signal (S_{Glu}) and dendritic spine's voltage (V), respectively. C_M denotes the membrane capacitance, R_M the membrane resistance, E_{Ca2+} the cell's Nernst potential for Ca^{2+}, and I(t) the external input.

I propose that in a synaptic single-compartment model with an NMDAR population, the glutamate and voltage gates can be considered as separate variable resistors, and each NMDAR

as a composition of the glutamate gate resistor in-series with the voltage gate resistor (**Figure 3.14**). Therefore, in a single synapse, the overall post-synaptic NMDAR conductance, g = g(t), can be expressed as the product of the NMDARs' gate conductances, divided by their sum:

$$g = \frac{g_{Glu} \cdot g_V}{g_{Glu} + g_V} \quad , \tag{3.18}$$

where $g_{Glu} = g_{Glu}(t)$ represents the conductance of the NMDAR population's glutamate gates, and $g_V = g_V(t)$ represents the conductance of its voltage gates. g determines the level of permeability and Ca^{2+} influx through channels of the NMDAR population (**Figure 3.14**). The product of g_{Glu} and g_V in **Eq. (3.18)** also implies necessity of coincidence of glutamate and voltage gate activation.

The voltage dependence of an NMDAR has been previously modeled, using experimental data, as a logistic function of synaptic Mg^{2+} concentration and the post-synaptic voltage (Dayan & Abbott, 2001; Jahr & Stevens, 1990). We can similarly describe the conductance of a population of NMDARs' voltage gates, g_V, as a simplified logistic function of the normalized voltage along the dendrite, $V_D = V_D(t)$:

$$g_V = \frac{1}{1 + e^{a_v \cdot V_D + b_v}} \quad , \tag{3.19}$$

with constants a_v and b_v provided in **Table 3.2**. In **Eq. (3.19)**, we assume that Mg^{2+} concentration stays relatively constant in the synapse, which is typical for physiological conditions.

There are few studies on modeling the dynamics of the NMDAR glutamate gate and how they change with different NMDAR subunit compositions. We are particularly interested in the glutamate gate characteristic activation time defined by the numbers of fast and slow NMDARs (Shouval et al., 2002). So, let τ_{Syn} be the characteristic time with which g_{Glu} approaches the

glutamate gate conductance limit, g_L. Evolution of g_{Glu} can be written in the form of a single-time-step mapping:

$$g_{Glu}^{t+1} = g_L + (g_{Glu}^{t} - g_L) \cdot e^{-\Delta t / \tau_{Syn}} \quad , \tag{3.20}$$

as derived from approximate solution of:

$$\frac{dg_{Glu}}{dt} = \frac{g_L - g_{Glu}}{\tau_{Syn}} \quad , \text{ where} \tag{3.21}$$

$$g_L = \frac{1}{1 + e^{a_L \cdot S_{Glu}(t) + b_L}} \quad , \tag{3.22}$$

Δt is a single time step, a_L and b_L are constants, the values for which are provided in **Table 3.2**. Here, I assume that glutamate gate conductance limit, g_L, depends on S_{Glu} (**Eq. (3.22)**), because the number of glutamate gates that are activated is limited by the synaptic glutamate concentration, and is constrained by the total number of NMDARs in the synapse.

NMDAR Characteristic Time

τ_{Syn} defines the rate at which g reaches its limit, g_L, and also depends on subunit composition of NMDARs, which can change over time. I assume τ_{Syn} can be expressed through the numbers of slow and fast NMDARs (n_{Slow} and n_{Fast}) and their time constants (τ_{Fast} and τ_{Slow}) as follows:

$$\tau_{Syn} = \frac{n_{Slow}\tau_{Slow} + n_{Fast}\tau_{Fast}}{n_{Total}}, \tag{3.23}$$

where n_{Total} is the total number of NMDAR receptors.

I propose *a key point and assumption of the SITDL hypothesis, which is that NMDAR subunit switching, and consequently, changes in τ_{Syn}, depend on the difference between the NMDAR gate conductances, g_{Glu} and g_V. Optimal Ca^{2+} influx occurs when the value of g_{Glu} matches that of g_V*. Let us define time evolution of τ_{Syn} as:

$$\tau_{Syn}^{t+1} = \tau_{Syn}^{t} + \gamma \cdot \Delta\tau \cdot (g_{Glu} - g_V)(g_L - g_{Glu}), \tag{3.24}$$

where $\Delta\tau$ is the step size for changing τ_{Syn}, γ is a scaling constant, and (g_{Glu} - g_V) is the gate conductance mismatch. The primary goal of the SITDL model is to find a τ_{Syn} that minimizes the gate conductance mismatch, which would effectively represent the estimated time difference between the synaptic glutamate signal and the dendritic voltage signal. **Eq. (3.24)** is equivalent to minimization of the function $F(\tau_{Syn}) = (g_{Glu} - g_V)^2$.

Note that using **Eq. (3.23)**, and the assumption that n_{Total} stays constant, numbers of slow and fast NMDARs can be calculated from τ_{Syn} as follows:

$$n_{Slow} = \frac{n_{Total}\left(\tau_{Syn} - \tau_{Fast}\right)}{\tau_{Slow} - \tau_{Fast}}, \tag{3.25}$$

$$n_{Fast} = n_{Total} - n_{Slow}. \tag{3.26}$$

Synaptic Current and Dendritic Voltage

We can then calculate the synaptic current, $I_{Syn} = I_{Syn}(t)$, that results from NMDAR activation and consequent Ca^{2+} influx, using overall NMDAR conductance, g, and dendritic voltage, V_D:

$$I_{Syn} = g \cdot V_D. \tag{3.27}$$

Changes in dendritic voltage are described via a single-compartment conductance-based model:

$$C\frac{dV_D}{dt} = \frac{V_{Rest} - V_D}{R} + I_D(t - \tau_D) + k \cdot I_{Syn}(t), \tag{3.28}$$

where, V_{Rest} is resting membrane potential, $\tau_R = CR$ is a time constant, k ($k \leq 1$) is a current contribution constant, $I_D(t - \tau_D)$ is delayed dendritic signal, and I_{Syn} is synaptic current resulting from NMDAR activation.

Stabilization of NMDAR Activation Time τ_{Syn}

The change in τ_{Syn} as described by **Eq. (3.24)** is always dominated by the gate conductance mismatch. Therefore, unless there is a consistent and perfect match of the gate conductances, g_{Glu} and g_V, τ_{Syn} will constantly change. I propose a stabilization mechanism that depends on the overall NMDAR conductance, g, such that when g is consistently large, τ_{Syn} will eventually stabilize. We modify **Eq. (3.24)** to include a stabilization variable, P, with the following set of equations:

$$\tau_{Syn}^{t+1} = \tau_{Syn}^{t} + P \cdot \Delta\tau_{Syn}, \tag{3.29}$$

where

$$\Delta\tau_{Syn} = \gamma \cdot \Delta\tau \cdot (g_{Glu} - g_V)(g_L - g_{Glu}), \tag{3.30}$$

$$P = \frac{1}{1 + e^{a_P \cdot \sigma^t + b_P}}, \tag{3.31}$$

$$\sigma^{t+1} = \sigma^t + \frac{g \cdot \left(\Delta\tau_{Max} - |\Delta\tau_{Syn}|\right)}{\tau_P}. \tag{3.32}$$

P and σ define the stabilization mechanism: when g is consistently greater than zero (**Eqs. (3.31), (3.32)**), and concurrently the change in τ_{Syn}, $\Delta\tau_{Syn}$, is close to zero, τ_{Syn} stops changing. $\Delta\tau_{Max}$ is

the maximum possible value for $\Delta\tau_{Syn}$. Its value and the values for constants a_P, b_P, and τ_P are provided in **Table 3.2**.

SITDL Multi-Synaptic Memory and Recall

A silent synapse, without AMPARs, may change and stabilize its NMDAR population. Once the synapse has matured, and AMPARs have been inserted into the membrane, this may provide the synapse with a way of "recalling" the memorized timing difference. In this case, the NMDAR population no longer has to rely on passing dendritic voltage signals for activation of their voltage gates, as AMPARs can provide depolarization in response to the glutamate signal. Thus, a glutamate signal alone could, in theory, provide activation of NMDARs' glutamate gates, and also indirectly provide depolarization for activation of NMDARs' voltage gates through the AMPARs. The resulting overall NMDAR current depends on $\tau_{Syn,}$ of the stabilized NMDAR population, which encoded the previously memorized timing difference, in effect, recalling it. In particular, this may be useful when there are many synapses, as there often are in a developing neuron. If these synapses are able to learn and memorize the timing differences, or commonly occurring inter-spike intervals, of recurring signals during development, then once they mature, they may be able to reproduce certain memorized parts or patterns from the original developmental signals.

Therefore, to test whether SITDL mechanisms could be used to memorize peak times of a glutamate signal and reconstruct it, several alterations are made. Multiple synapses are used, with each synapse represented as a single-compartment SITDL model with stabilization mechanisms (**Eqs. (3.12)-(3.32)**, with its own dendritic delay time constant, τ_D, and initial glutamate gate activation time, τ_{Syn}. A set of "Learning Phase" simulations starts with a full set of synapses (i =

1, …,N), with τ_D values uniformly distributed between 100 ms and 1300 ms, and initial τ_{Syn} = 50 ms. In these simulations, each synapse receives a periodic glutamate signal as described before, and a sparse voltage signal with one spike repeating over the same period as the glutamate signal (**Figure 3.20B**). For each synapse *i*, g_{Avg}, the overall NMDA conductance averaged over the last fifth of the simulation time, is calculated. If a synapse's g_{Avg} is lower than a threshold factor, δ, times the value of g_{Avg} averaged across the entire set of synapses, such that $g^i_{Avg} < \frac{\delta}{N}\sum_{j=1}^{N} g^j_{Avg}$, then it is eliminated. All synapses that have not been eliminated are considered stabilized, with τ_{Syn} permanently fixed. I assume that stabilized synapses express AMPARs, which depolarize the synapse upon binding glutamate. Therefore, for "Recall Phase" simulations, each stabilized synapse receives a single glutamate spike coincident with a single voltage spike (**Figure 3.20C**). Note that k, the constant for synaptic current contribution to voltage is set to 1.0 for these simulations. The resulting synaptic voltage traces are shifted by their corresponding τ_D and summed over all stabilized synapses to give a "recall signal", which is compared against the original glutamate signal used during Learning Phase simulations. "No-SITDL" simulations use the same Learning Phase and Recall Phase simulations, except there are no SITDL mechanisms during the Learning Phase, such that there are no changes in τ_{Syn} for any synapse.

Mutual Information (MI) Analysis

A mutual information (MI) estimator, Adaptive partition using Interspike intervals MI Estimator (AIMIE), was used to compare recall signal and original glutamate signal (Ekaterina Dmitrievna Gribkova et al., 2018). Short signals were repeated with proper periodicity, such that each signal had at least 4000 spikes for more accurate MI estimation. A higher MI estimate for the two

signals suggests that they have a greater dependency and higher degree of similarity. AIMIE has been shown to work well with spike time series that have disparate firing rates, and seems to provide more accurate estimates of MI than several other commonly used MI estimators. Details on AIMIE's calculations and its use in information flow analysis of a spiking network are provided in **Chapter 4**.

Table 3.2. SITDL Model Constants

Symbol	Description	Value *(no units, unless indicated)*
σ_N	Dispersion constant for S_N	10.0
α_L	Exponential constant for S_L	0.10
I_0	Constant current for I_{Sig} and I_{Syn}	0.04
a_V	Exponential constant for $g_V(t)$	-8.0
b_V	Exponential constant for $g_V(t)$	6.0
a_L	Exponential constant for $g_L(t)$	-10.0
b_L	Exponential constant for $g_L(t)$	3.0
Δt	Simulation time step constant	0.01 ms
τ_{Fast}	Time constant for glutamate gate activation of fast NMDARs	7.0 ms
τ_{Slow}	Time constant for glutamate gate activation of slow NMDARs	50.0 ms
n_{Total}	Total number of NMDAR receptors	50 receptors
γ	Constant for scaling the time step $\Delta\tau$	1.0
$\Delta\tau$	Time step for changing τ_{Syn}	0.008 ms
V_{Rest}	Normalized resting potential	0.0
τ_R	Time constant for changes in V(t)	1.0 ms
k	Constant for current contribution	0.40
τ_P	Time constant for determining how slowly $\sigma(t)$ changes	0.25 ms
$\Delta\tau_{Max}$	Maximum of $\Delta\tau_{Syn}$, used for modulating changes in $\sigma(t)$	0.02 ms
a_P	Exponential constant for logistic function of $\sigma(t)$ in P(t)	0.30
b_P	Exponential constant for logistic function of $\sigma(t)$ in P(t)	-70.0
δ	Threshold factor for synaptic elimination	1.30

All simulations of the SITDL model were programmed and run in Spyder 3.0.0, a Python 3.5 environment. SITDL data was analyzed using MATLAB R2016b and Excel 2016.

RESULTS:

Changes in Synaptic Timing in the SITDL Model

The core mechanism of the SITDL model involves learning the timing difference between synaptic glutamate and dendritic voltage involving changes in NMDAR glutamate gate activation time, τ_{Syn}, and correspondingly, numbers of fast and slow NMDARs. Using periodic glutamate and voltage signals and computational single-compartment model described in Methods (**Eqs. (3.12)-(3.28)**), I ran multiple simulations for a range of initial values of dendritic delay, τ_D, and τ_{Syn}.

Figure 3.15 shows evolution of τ_{Syn} and NMDAR conductances for a single synapse receiving glutamate signal, S_{Glu}, followed by dendritic voltage signal, V_D. Both are copies of a similar periodic signal, with dendritic signal delayed by time τ_D. The timing difference between signals produces a gate conductance mismatch, (g_{Glu} - g_V) (**Eq. (3.24)**), changing τ_{Syn} and shifting g_{Glu} peaks. This leads to greater overlap of the gate conductances. With initial values τ_{Syn} = 1 ms and τ_D = 30 ms (**Figure 3.15A**), glutamate gate conductance, g_{Glu}, initially peaks before voltage gate conductance, g_V. This results in positive gate conductance mismatch and increase of τ_{Syn}, during the rise of g_{Glu}. Coincidence appears to have been achieved at the end of the 2000 ms simulation. **Figure 3.15B** shows similar coincidence with initial values τ_{Syn} = 10 ms and τ_D = 30 ms. In **Figure 3.15C**, g_V initially peaks before g_{Glu}, and τ_{Syn} decreases over time, resulting in greater overlap of gate conductances.

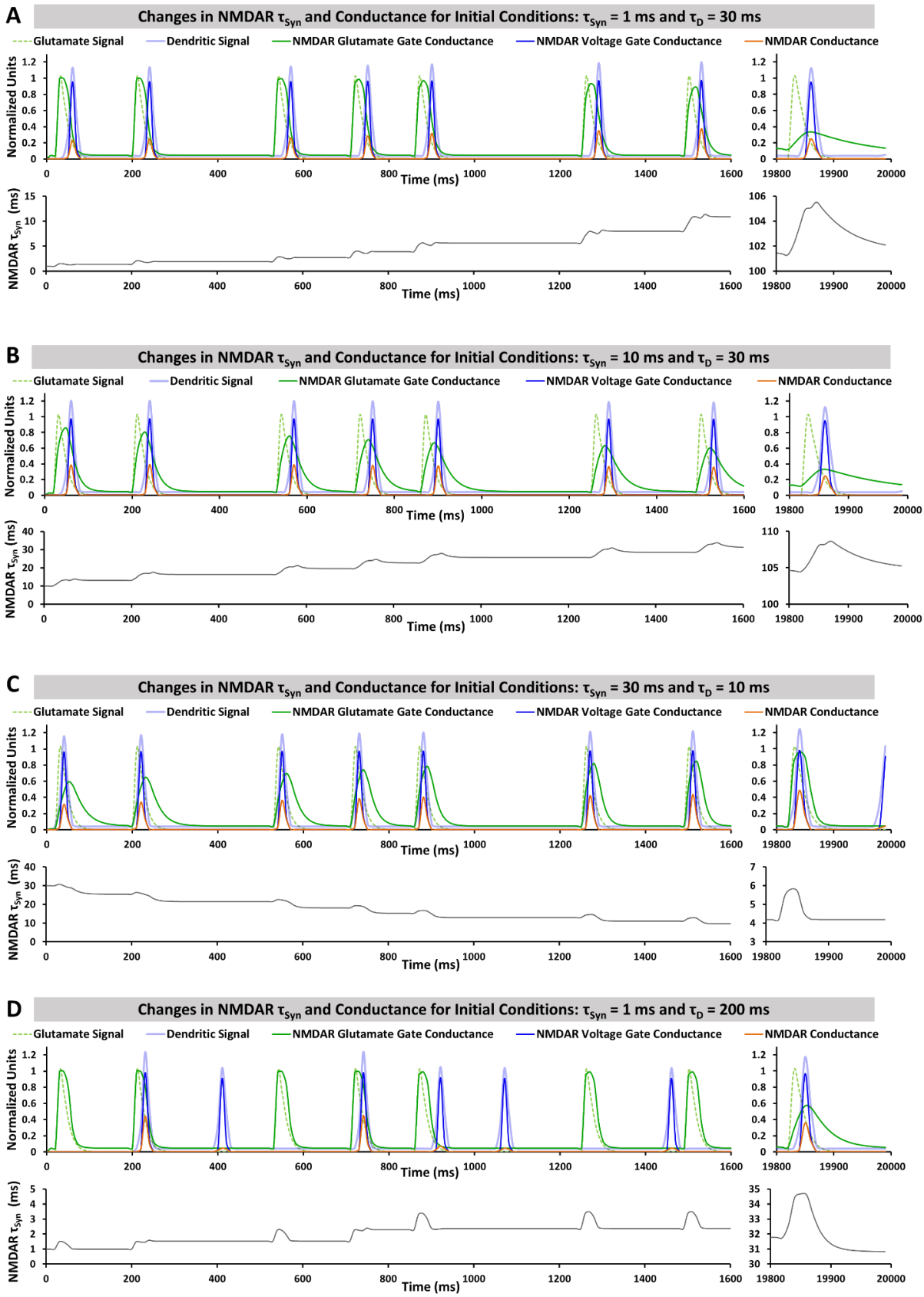
A
Changes in NMDAR τ_{Syn} and Conductance for Initial Conditions: τ_{Syn} = 1 ms and τ_D = 30 ms
B
Changes in NMDAR τ_{Syn} and Conductance for Initial Conditions: τ_{Syn} = 10 ms and τ_D = 30 ms
C
Changes in NMDAR τ_{Syn} and Conductance for Initial Conditions: τ_{Syn} = 30 ms and τ_D = 10 ms
D
Changes in NMDAR τ_{Syn} and Conductance for Initial Conditions: τ_{Syn} = 1 ms and τ_D = 200 ms
Glutamate Signal
Dendritic Signal
NMDAR Glutamate Gate Conductance
NMDAR Voltage Gate Conductance
NMDAR Conductance
Normalized Units
NMDAR τ_{Syn} (ms)
Time (ms)

Figure 3.15

Figure 3.15. SITDL Simulations of a single synapse receiving periodic inputs and showing evolution of τ_{Syn} and conductances for several different initial conditions of τ_{Syn} and τ_D. In each case, the synapse receives a synaptic glutamate signal (dashed light green) followed by a dendritic voltage signal (light blue) delayed by a time τ_D. τ_{Syn} changes over time due to the NMDAR gate conductance mismatch. **A)** Simulation with initial values τ_{Syn} = 1 ms and τ_D = 30 ms. Initially, glutamate gate conductance, g_{Glu}, peaks before voltage gate conductance, g_V. Over time, τ_{Syn} grows, delaying the g_{Glu} peak and achieving greater coincidence. **B)** Simulation with initial values τ_{Syn} = 10 ms and τ_D = 30 ms. τ_{Syn} grows over time, and achieves similar coincidence to A. **C)** Simulation with initial values τ_{Syn} = 30 ms and τ_D = 10 ms. In this case, g_V initially peaks before g_{Glu}. τ_{Syn} decreases over time, resulting in greater coincidence of the gate conductances. **D)** Simulation with initial values τ_{Syn} = 1 ms and τ_D = 200 ms. Though there is much less overlap between glutamate and voltage signals, due to periodicity of the signal, the SITDL mechanism still appears to achieve greater coincidence over time.

If initial value τ_{Syn} = 1 ms and τ_D is set to much larger value (τ_D = 200 ms, **Figure 3.15D**), then there will be significantly less overlap between the glutamate and voltage signal spikes. Because there is still some overlap and the signals are periodic, τ_{Syn} grows slowly and still appears to achieve coincidence by the end of the simulation.

The examples in **Figure 3.15A-C** may relate to delay line systems of avian and mammalian auditory brainstem and of weakly electric fish (Ashida & Carr, 2011; Carr, 1986), where a single neuron can receive multiple copies of the same or similar pattern with small delays. Mechanisms of SITDL model may be particularly useful in establishing the precise timing in developing synapses of delay line systems, which have been shown to involve NMDARs (Tang & Carr, 2007). As seen in **Figure 3.15D**, SITDL mechanisms can still work to achieve greater gate conductance coincidence even when dendritic delay, τ_D, is significant and there is much less overlap in signals.

Figure 3.16 demonstrates simulations of the SITDL model with a greater range of initial delays to explore the evolution of τ_{Syn}.

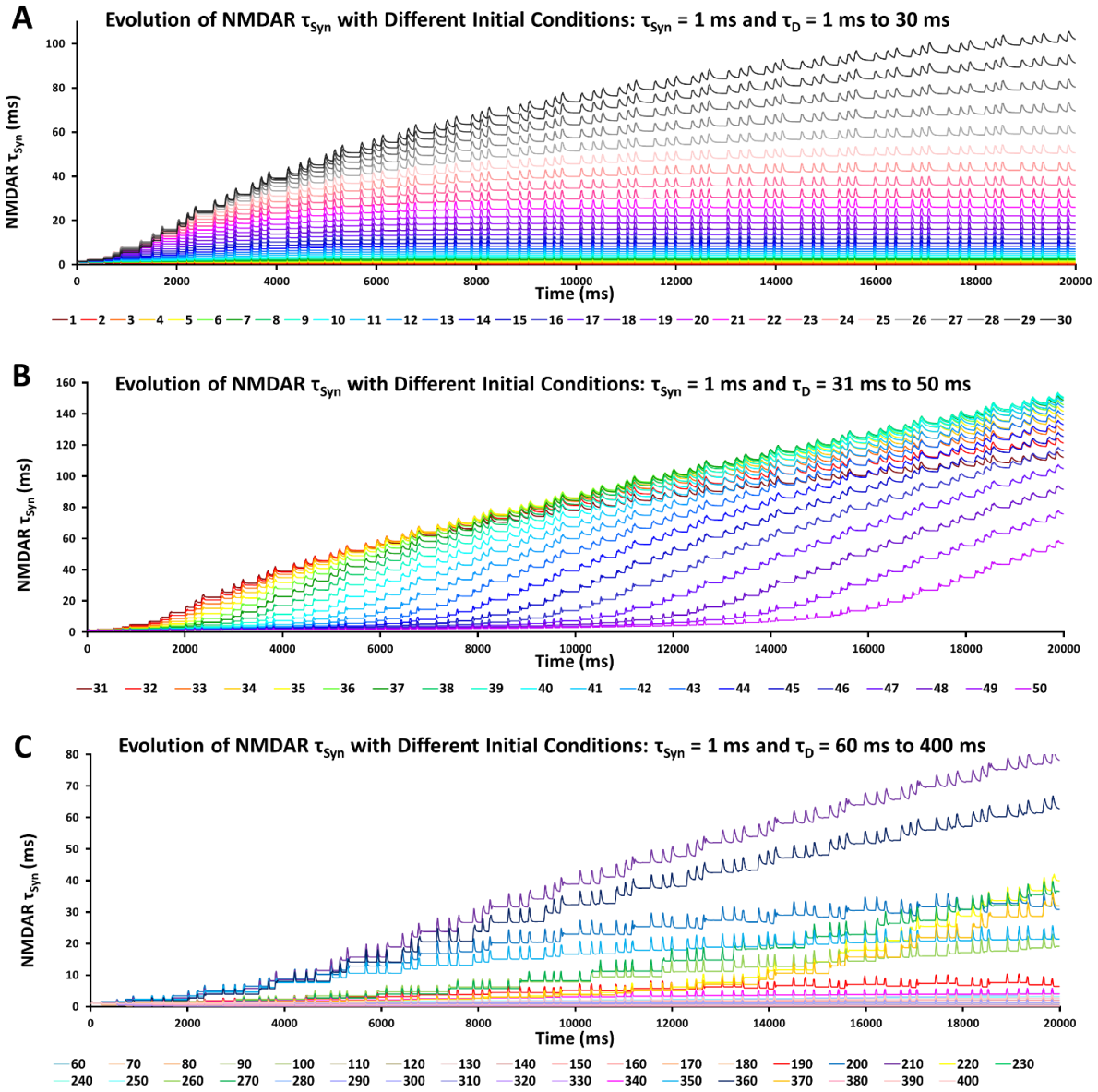

Figure 3.16. Simulations of SITDL model with initial conditions τ_{Syn} = 1 ms and τ_D ranging from 1 ms to 400 ms. Each curve shows the changes in τ_{Syn} over time, for a single simulation with the specific initial τ_D value indicated by the color of the curve. **A)** Changes in τ_{Syn} when starting with τ_{Syn} = 1 ms and different initial τ_D value for each curve, ranging from 1 ms to 30 ms. By 20000 ms simulation time, each curve appears to reach more stable values, particularly for τ_D = 25 ms and below. **B)** Changes in τ_{Syn} with initial conditions τ_{Syn} = 1 ms and τ_D value ranging from 31 ms to 50 ms. In these cases, τ_{Syn} increases, but there does not appear to be sufficient simulation time for it to reach more stable values as in A. **C)** Changes in τ_{Syn} with initial conditions τ_{Syn} = 1 ms and τ_D value ranging from 60 ms to 400 ms, which provide much less overlap of glutamate and voltage signals. Certain τ_D values, such as 190 to 220 ms, 260 to 270 ms, and 340 to 370 ms, seem to provide sufficient overlap for τ_{Syn} to change significantly and potentially achieve coincidence of gate conductances with enough simulation time.

For initial τ_{Syn} = 1 ms, and τ_D = 1 ms to 25 ms, τ_{Syn} seems to quickly attain relatively stable values, with very little change towards the end of the simulation. The smaller the τ_D, the more quickly τ_{Syn} stabilizes. This can be explained by smaller gate conductance mismatch and greater overlap between glutamate and voltage signals. For τ_D =31 to 50 ms (**Figure 3.16B**), τ_{Syn} grows, but appears to change more slowly and does not stabilize as much as for smaller τ_D . Longer simulation times may allow τ_{Syn} to stabilize for these larger τ_D values. For τ_D = 60 to 400 ms (**Figure 3.16C**), in some cases there is insufficient overlap of signals to cause changes in τ_{Syn}. Certain τ_D values, such as τ_D = 190 to 220 ms, 260 to 270 ms, and 340 to 370 ms, seem to provide sufficient overlap for τ_{Syn} to change significantly and potentially achieve coincidence of glutamate and voltage gate conductances with enough simulation time. In particular, for τ_D = 190, 200, and 350 ms, τ_{Syn} appears to approach stable values at the end of 20000 ms simulation time.

Note that while τ_{Syn} approaches a more stable range of values in most cases, it does not fully stabilize. Therefore I made key alterations to the SITDL model, using **Eqs. (3.29)-(3.32)** for implementing τ_{Syn} stabilization mechanisms, which are important for demonstrating memory formation and recall. **Figure 3.17** demonstrates simulations of this variation of SITDL model for τ_D ranging from 1 to 400 ms, similar to **Figure 3.16**, with initial τ_{Syn} of 100 ms and longer simulation time (40000 ms). This shows, τ_{Syn} can decrease from a high initial value (τ_{Syn} = 100 ms), and fully stabilize at values similar to those in **Figure 3.16**, due to SITDL stabilization mechanism.

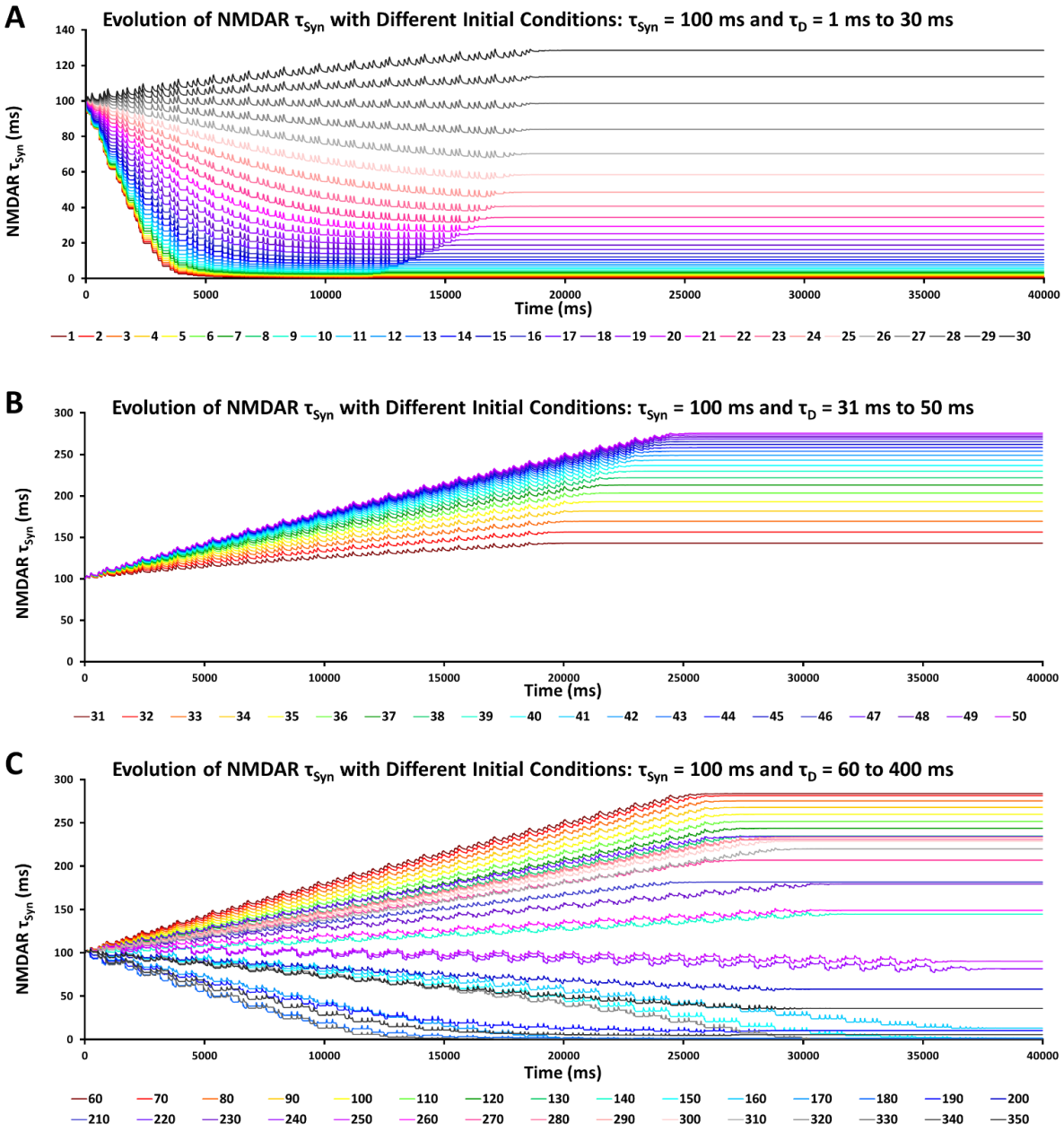

Figure 3.17. Simulations of SITDL model with stabilization mechanism for τ_{Syn}, and with initial conditions τ_{Syn} = 100 ms and τ_D ranging from 1 ms to 400 ms. Each curve shows the changes in τ_{Syn} over time, for a single simulation of 40000 ms, with the specific initial τ_D value indicated by the color of the curve. The figure is very similar to **Figure 3.16**, however, it shows that τ_{Syn} can also decrease from a high initial value, and can stabilize fully. **A)** Changes in τ_{Syn} when starting with τ_{Syn} = 100 ms and different initial τ_D value for each curve, ranging from 1 ms to 30 ms. τ_{Syn} stabilization occurs for all τ_D, and is faster for smaller τ_D. **B)** Changes in τ_{Syn} with initial conditions τ_{Syn} = 100 ms and τ_D value ranging from 31 ms to 50 ms. **C)** Changes in τ_{Syn} with initial conditions τ_{Syn} = 100 ms and τ_D value ranging from 60 ms to 400 ms. Note that even with very limited overlap of the periodic signals, τ_{Syn} still stabilizes for most values at the end of the simulation.

Slow and Fast NMDARs in the SITDL Model

I assumed that τ_{Syn} is determined by the numbers of fast and slow NMDARs. Fast NMDARs can be considered those with GluN1/GluN2A subunit composition, with estimated rise-to-peak time of 7 ms, and slow NMDARs as those with GluN1/GluN2B subunit composition, with estimated rise-to-peak time of 50 ms (Singh et al., 2011).

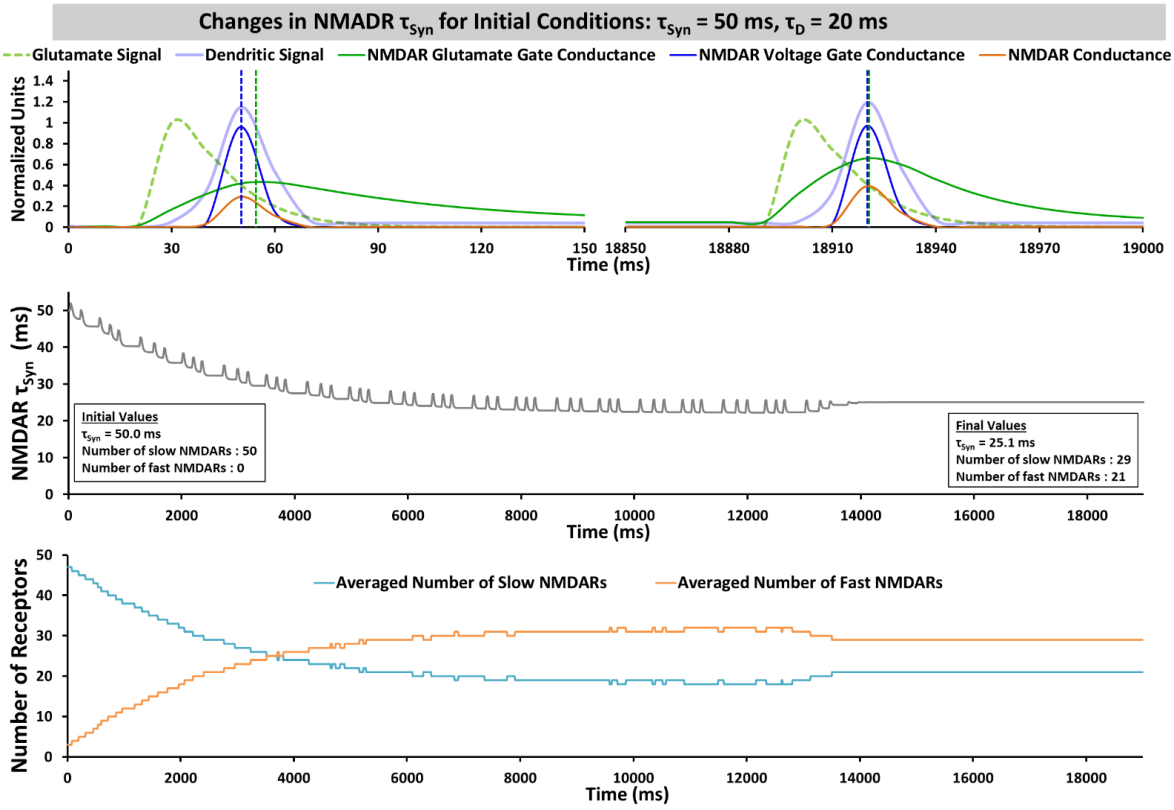

Figure 3.18. Simulations of SITDL synapse model with plasticity and stabilization mechanisms for τ_{Syn}, and with initial conditions τ_{Syn} = 50 ms and τ_D = 20 ms, and constants for calculating the numbers of slow and fast NMDARs n_{Total} = 50, τ_{Fast} = 7 ms, and τ_{Slow} = 50 ms. Green and blue dashed vertical lines respectively denote the peaks of glutamate gate conductance, g_{Glu} and voltage gate conductance, g_V. The top graphs show the first 150 ms of the simulation at the left, and last 150 ms at the right, with g_{Glu} and g_V peaks noticeably becoming more coincident. The middle graph shows the time course of τ_{Syn}, with an initial value of 50 ms, which corresponds to the synapse having 50 slow NMDARs and 0 fast NMDARs. τ_{Syn} decreases overall, with some oscillation, until it stabilizes at 25.1 ms, which corresponds to 21 slow NMDARs and 29 fast NMDARs in the synapse, indicating a replacement of slow NMDARs with fast NMDARs. The bottom plot show the corresponding numbers of slow and fast NDMARs over time, each averaged over a 480 ms time window.

Over the course of development, slow NMDARs are typically replaced with fast NMDARs (Paoletti et al., 2013). As a demonstration, I ran a SITDL model simulation with initial values τ_{Syn} = 50 ms and τ_D = 20 ms, and calculated numbers of fast and slow NMDARs using **Eqs. (3.25)-(3.26)** with n_{Total} = 50, τ_{Fast} = 7 ms, and τ_{Slow} = 50 ms.

Due to decreasing τ_{Syn}, g_{Glu} maximum shifts significantly closer to g_V maximum (**Figure 3.18**, top). Initially, τ_{Syn} = 50 ms, with 50 slow NMDARs and 12 fast NMDARs. As τ_{Syn} decreases and stabilizes at its final value of ~25 ms, there are 21 slow NMDARs and 29 fast NMDARs (**Figure 3.18**, middle and bottom plots). Reduction of τ_{Syn} can therefore represent a replacement of slow NMDARs with fast NMDARs, and τ_{Syn} convergence can be considered as stabilization in expression of fast and slow NMDARs.

SITDL Multi-Synaptic Memory Recall and MI Analysis

The potential for SITDL mechanisms to achieve greater coincidence between NMDAR glutamate and voltage gate conductances, even with very limited overlap of signals, suggests that they could be used to memorize peak times of a glutamate signal and reconstruct it. To explore this, large numbers of SITDL synapses, with stabilization mechanisms and τ_D values uniformly distributed between 100 ms and 1300 ms, were run through Learning Phase simulations, quite similar to previous simulation with ranges of different initial conditions. During the Learning Phase simulation, each synapse receives a periodic glutamate signal, and a sparse rhythmic voltage signal of the same period (Kamioka et al., 1996; Lippe, 1994), delayed relative to the glutamate signal by time constant, τ_D (**Figure 3.20B**). We can consider this initial set of synapses to be located along a neuron's dendrite, receiving the same glutamate signal from the branched axon of a single pre-synaptic neuron (**Figure 3.19**). These types of redundant multi-synaptic

connections are often seen in neocortex, hippocampus, in neural development, and they may also arise as a result of LTP (Bartol Jr et al., 2015; Hiratani & Fukai, 2018; Kasthuri et al., 2015; Toni et al., 1999).

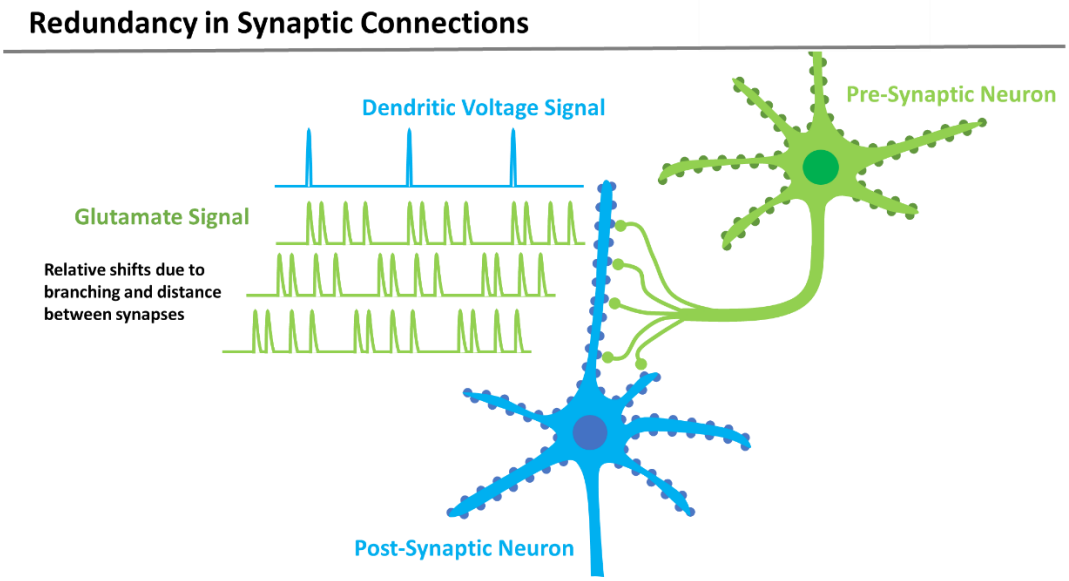

Figure 3.19. Redundancy in synaptic connections can cause a range of relative shifts between synaptic glutamate signal and dendritic voltage signal, due to differences in axonal branching and distances between synapses.

As before, τ_{Syn} and conductance values change over the course of the Learning Phase simulations (**Figure 3.20**), and after simulation end, synapses with insufficient overall NMDAR conductance are eliminated, similar to synaptic pruning. **Figure 3.20A** shows the changes in τ_{Syn} for synapses that were stabilized, with an initial set of synapses having initial conditions τ_{Syn} = 50 ms, and τ_D uniformly distributed from 100 ms to 1300 ms, with a step size of 10 ms. The τ_D step can represents how far apart the initial synapses are on the dendrite (**Figure 3.19**). The leftover stabilized synapses are assumed to express AMPARs. These were run through Recall Phase simulations, where each stabilized synapse receives a single glutamate spike and a single voltage spike that coincide due to AMPAR expression (**Figure 3.20C**).

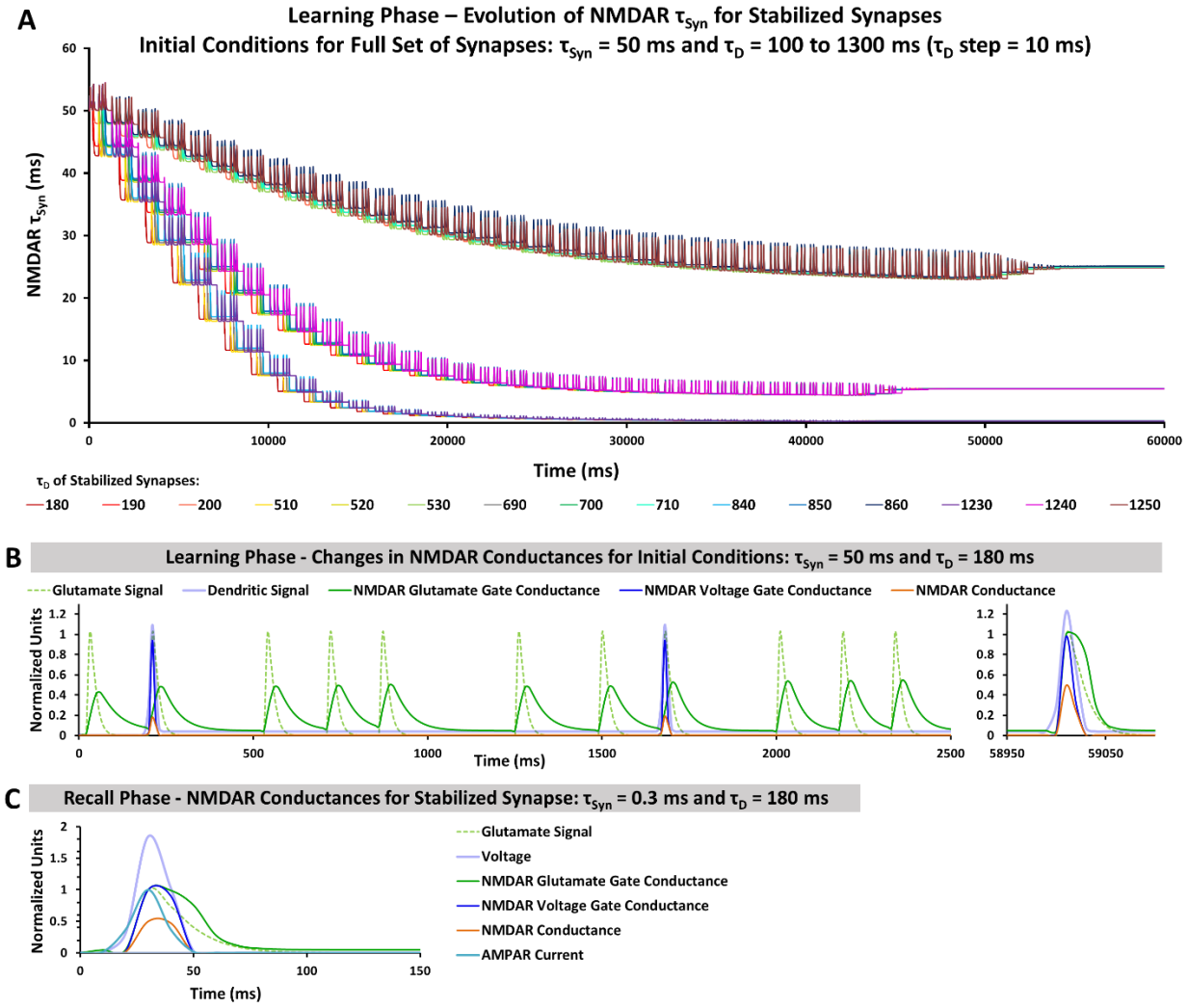

Figure 3.20. Multi-synaptic SITDL simulations showing changes in τ_{Syn} and conductances for stabilized synapses. Learning Phase simulations start with a full set of synapses, with initial conditions τ_{Syn} = 50 ms and τ_D values uniformly distributed from 100 ms to 1300 ms, with a step of 10 ms. Each synapse receives periodic glutamate and voltage signals, with a relative delay of τ_D, as shown in B. At the end of the simulations, synapses with insufficient average overall NMDAR conductance are eliminated, leaving only stabilized synapses. **A)** Evolution of τ_{Syn} for synapses that were stabilized at the end of the simulation. **B)** An example of a Learning Phase simulation, showing how NMDAR conductances change over time for a synapse with initial τ_{Syn} = 50 ms and τ_D = 180 ms. The synapse receives a periodic glutamate signal and a sparse voltage signal with the same period. The first voltage spike occurs close to the second spike of the glutamate signal. At the end of the simulation the average overall NMDAR conductance is sufficient for this synapse to be stabilized. **C)** Recall Phase simulation of stabilized synapse from B with final values τ_{Syn} = 0.3 ms and τ_D = 180 ms. Since stabilized synapses are assumed to express AMPARS, in this case the synapse receives a single glutamate spike coincident with a single voltage spike.

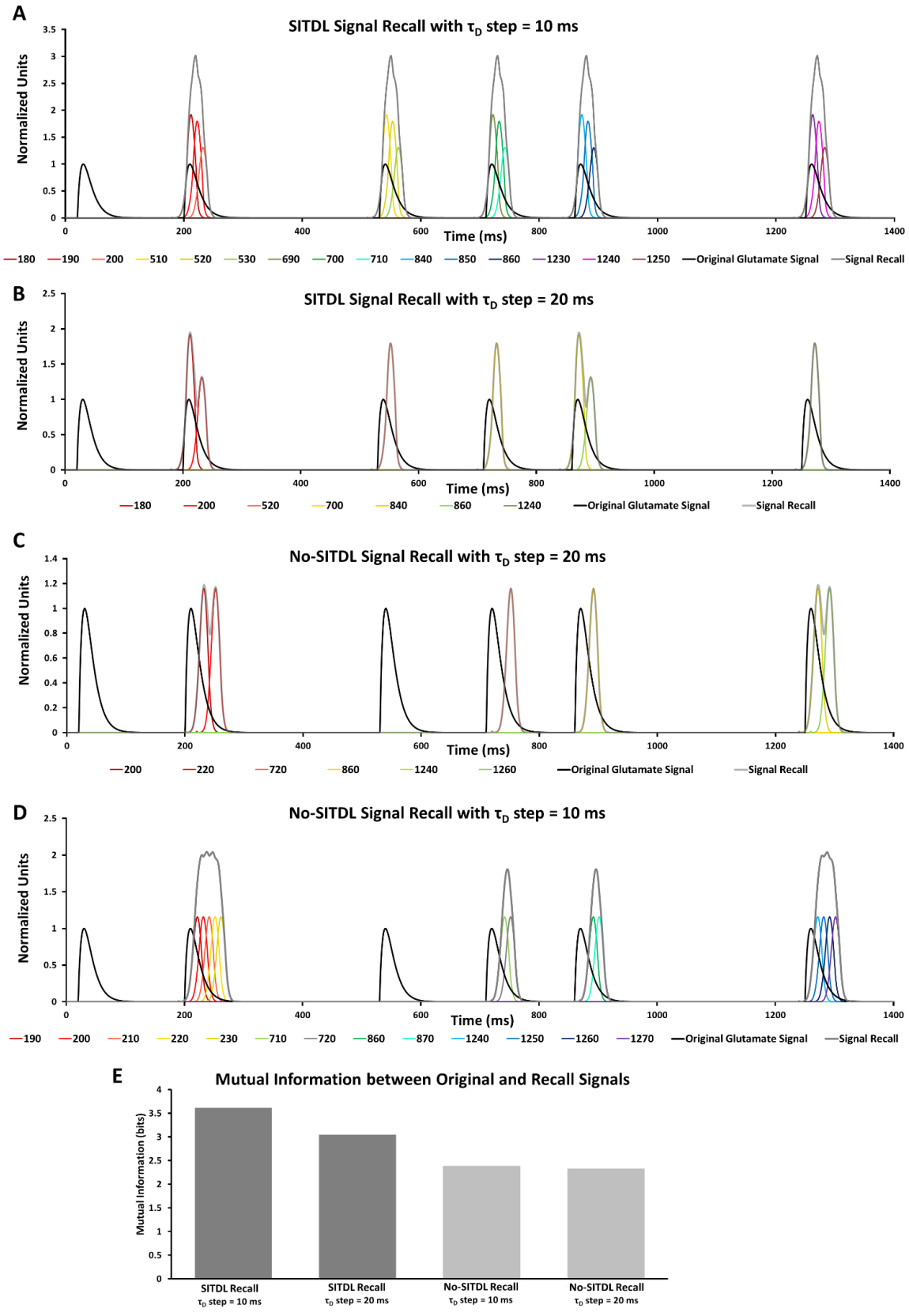
A
SITDL Signal Recall with τ_D step = 10 ms
Normalized Units
Time (ms)
180 190 200 510 520 530 690 700 710 840 850 860 1230 1240 1250 Original Glutamate Signal Signal Recall
B
SITDL Signal Recall with τ_D step = 20 ms
Normalized Units
Time (ms)
180 200 520 700 840 860 1240 Original Glutamate Signal Signal Recall
C
No-SITDL Signal Recall with τ_D step = 20 ms
Normalized Units
Time (ms)
200 220 720 860 1240 1260 Original Glutamate Signal Signal Recall
D
No-SITDL Signal Recall with τ_D step = 10 ms
Normalized Units
Time (ms)
190 200 210 220 230 710 720 860 870 1240 1250 1260 1270 Original Glutamate Signal Signal Recall
E
Mutual Information between Original and Recall Signals
Mutual Information (bits)
SITDL Recall τ_D step = 10 ms
SITDL Recall τ_D step = 20 ms
No-SITDL Recall τ_D step = 10 ms
No-SITDL Recall τ_D step = 20 ms

Figure 3.21

Figure 3.21. Results of Recall Phase simulations for synapses with and without SITDL mechanisms. There are 4 different sets of synaptic simulations, each with their own starting conditions: 2 sets of SITDL simulations with τ_D step = 10 ms and 20 ms, and 2 sets of simulations without SITDL mechanisms with τ_D step = 10 ms and 20 ms. For all simulation sets, τ_D values are initially distributed uniformly with the corresponding τ_D step, and all initial τ_{Syn} = 50 ms. For each simulation set, following Learning Phase simulations, synaptic elimination, and Recall Phase simulations, a recall signal is obtained by summating the resulting Recall Phase voltage signals (**Figure 3.20C**), shifted by corresponding τ_D (A-D traces in color), over all stabilized synapses. **A)** Recall signal for SITDL simulation set with τ_D step = 10 ms very closely resembles the original glutamate signal that was used during Learning Phase. **B)** Recall signal for SITDL simulation set with τ_D step = 20 ms is less similar to the original glutamate signal, because of the larger τ_D step used during Learning Phase. **C, D)** Recall signals for no-SITDL simulation sets with τ_D step = 10 ms and 20 ms, respectively. There are no changes in τ_{Syn} in these simulations, thus synapses are unable to achieve greater overlap of glutamate and voltage gate conductances. Even with synapses that were stabilized, the peaks of the recall signal are significantly shifted from the peaks of the original glutamate signal. **E)** Estimates of MI between original glutamate signal and recall signal for each simulation set, using MI estimator AIMIE (Ekaterina Dmitrievna Gribkova et al., 2018). SITDL simulation sets, particularly the one with smaller τ_D step (10 ms), provide greater MI estimates than no-SITDL simulation sets, suggesting greater degree of similarity between their reconstructed recall signal and the original glutamate signal.

The summation of resulting Recall Phase voltage signals across the stabilized synapses (**Figure 3.21**, traces in color), shifted by the corresponding τ_D value, provides a "recall signal" that very closely resembles the original glutamate signal from the Learning Phase (**Figure 3.21A**). For recall signals constructed from SITDL simulations with larger τ_D step, the peak times are less similar to those of the glutamate signal (**Figure 3.21B**). Furthermore, recall signals constructed from simulation sets (Learning Phase, elimination, Recall Phase) without SITDL mechanisms, such that τ_{Syn} cannot change, appear to be even less similar to the original signal, as they are missing the third spike of the glutamate signal (**Figure 3.21C-D**). Using MI estimator AIMIE (Ekaterina Dmitrievna Gribkova et al., 2018), I show that there seems to be a greater degree of similarity between the original signal and recall signal reconstructed from SITDL simulations (**Figure 3.21E**), showing that SITDL mechanisms across a set of synapses,

particularly when they are distributed closely enough, can be used to effectively memorize and recall synaptic signals.

DISCUSSION:

Using a computationally simple SITDL model, I show that a single synapse can "learn" the timing difference between glutamate and dendritic voltage signals. *The functioning principle of single-compartment SITDL model is to minimize NMDAR gate conductance mismatch, through switching of fast and slow NMDARs, and consequent changes in glutamate gate activation time, τ_{Syn}*. Optimal Ca^{2+} influx occurs when gate conductance mismatch is minimized, promoting stabilization of τ_{Syn}, and fast and slow NMDAR expression. Decreases in τ_{Syn} correspond to a replacement of slow NMDARs by fast NMDARs (**Figure 3.18**), which is typical during neural development. While the assumptions of the SITDL model have not yet been tested in physiological experiments, they may significantly expand synaptic capabilities in both biological and artificial systems. Much is still unknown about synaptic dynamics, NMDAR properties, and their activity-dependent modifications. It is possible that specific changes in NMDAR conformation, which occur independently for glutamate and voltage gate activation, may somehow convey the mismatch between gate conductances. Furthermore, SITDL provides potential mechanisms for new learning rules in ANNs, which have mostly employed synaptic weights changes. To our knowledge, SITDL is the first computational model that explores the ability of a single synapse to learn the timing of the signals it receives through changes in its receptor dynamics.

The SITDL model is able to increase overlap of g_{Glu} and g_V in a synapse that receives very similar periodic signals with small relative time shift (**Figure 3.15**). This may be

particularly useful for developing synapses in delay line systems of mammalian and avian auditory brainstem and sensory circuits of weakly electric fish (Ashida & Carr, 2011; Carr, 1986), all of which rely on very precise timing of signals. Some delay line systems resolve temporal disparities in the microsecond range even with just a single neuron (Kawasaki et al., 1988), and some depend on NMDARs in development (Tang & Carr, 2007). Notably, there have been few computational models for achieving precise timing in delay line systems. One such model explores temporal precision in the barn owl auditory system (Gerstner et al., 1996), which uses unsupervised Hebbian learning rules and a broad random distribution of transmission delays. While this may show how delay line systems with large numbers of neurons are able to achieve such temporal precision, it does not explain it can be achieved with much more limited numbers of neurons and transmission delays, such as in pre-pacemaker nucleus neurons of weakly electric fish (Kawasaki et al., 1988). Even without τ_{Syn} stabilization mechanisms, the SITDL model can still increase overlap of NMDAR gate conductances with larger dendritic delays, such as for some τ_D greater than 100 ms, which provide very little overlap of glutamate and voltage signals (**Figure 3.15D** and **Figure 3.16**). Due to the periodicity of the signals, less frequent but repeated overlap can cause significant changes in τ_{Syn}.

With τ_{Syn} stabilization mechanisms depending on overall changes in τ_{Syn}, and overall NMDAR conductance, the SITDL model has been shown to stabilize τ_{Syn} even with little overlap between glutamate and voltage signals (**Figure 3.17**). Changes in τ_{Syn} are bidirectional (**Figure 3.17**), and can represent replacement of fast and slow NMDARs (**Figure 3.18**). This relates to mechanisms of rapid bidirectional switching in NMDAR subunit compositions of developing hippocampus (Bellone & Nicoll, 2007).

Achieving greater gate conductance overlap with very limited overlap of signals may be quite useful, especially in memory formation. For instance, in a neuron that receives highly delayed copies of the same or similarly periodic signal, any memorized timing differences may represent the commonly occurring inter-spike intervals of the signal. I show that running SITDL simulations for a set of synapses receiving periodic signals with uniformly distributed τ_D, followed by synaptic elimination or stabilization based on average overall NMDAR conductance, it is possible to memorize and reconstruct a signal that is very similar, or practically the same as the original synaptic glutamate signal (**Figure 3.20-Figure 3.21**).

Potential Tests of SITDL in Biological Systems

Furthermore, if we consider the auto-associative networks of CA3 of the hippocampus, where synapses can exhibit bidirectional changes in NMDAR subunit compositions (Hunt et al., 2013), then the SITDL mechanism could potentially be used for sequence memorization in single neurons. **Figure 3.22** shows potential mechanism of sequence learning in CA3, with each synapse learning the timing between two distinct inputs. To explore the existence of SITDL mechanisms in physiological neurons, we may consider *in vitro* studies of neuronal cultures and brain slices, such as those of hippocampus and developing neocortex. In particular, it may be interesting to observe axonal and synaptic activity across multiple points in developing multi-synaptically connected neurons, possibly through multi-electrode recordings or calcium and voltage indicators. Furthermore, if neuronal activity can be manipulated, pharmacologically or with stimulating electrodes, alongside single particle tracking of different NMDAR subunits in post-synaptic membranes (Dupuis & Groc, 2020), then this may provide more insight into possible existence of neuronal SITDL mechanisms.

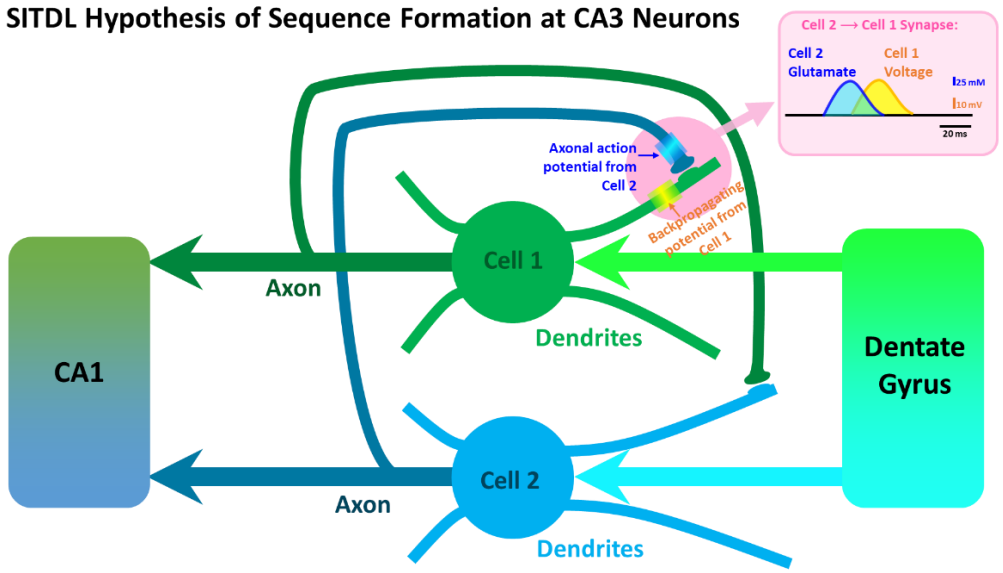

Figure 3.22. A hypothesis of how SITDL may enable memorization of sequences, as in place cell sequence formation, at CA3 neurons. Specifically, Cell 2 receives input from the dentate gyrus and fires, releasing glutamate onto Cell 1. In the meantime, Cell 1 also fires due to input from the dentate gyrus, and due to backpropagation of the potential along the dendrites of Cell 1, this results in NMDARs at the Cell 2 to Cell 1 synapse receiving a glutamate signal from Cell 2, followed by a backpropagating voltage signal from Cell 1. If SITDL mechanisms are in place, then the synapse may potentially learn the time difference between Cell 2 and Cell 1 activation. An entire CA3 network with these kinds of connections and mechanisms may therefore be capable of learning very complex sequences of activation.

CHAPTER 4: SUPPLEMENTARY MATERIAL

The following material provides details on the novel mutual information estimator, AIMIE, used for analysis in the previous chapter.

Mutual Information Estimator, AIMIE[2]

In this study, a novel mutual information estimator was developed to analyze information flow in a model thalamocortical network. My findings suggest that this estimator is a suitable tool for signal transmission analysis, particularly in neural circuits with disparate firing rates, and that the thalamic reticular nucleus can potentiate ascending sensory signals, while thalamic recipient cells in the cortex can recover mutual information in ascending sensory signals that is lost due to thalamic bursting.

The impact of thalamic state on information transmission to the cortex remains poorly understood. This limitation exists due to the rich dynamics displayed by thalamocortical networks and because of inadequate tools to characterize those dynamics. Here, we introduce a novel estimator of mutual information and use it to determine the impact of a computational model of thalamic state on information transmission. Using several criteria, this novel estimator, which uses an adaptive partition, is shown to be superior to other mutual information estimators with uniform partitions when used to analyze simulated spike train data with different mean spike rates, as well as electrophysiological data from simultaneously recorded neurons. When applied to a thalamocortical model, the estimator revealed that thalamocortical cell T-type

[2] *Journal of Neuroscience Permissions Policies:* https://www.jneurosci.org/content/rights-permissions. **Chapter contains previpusly published material from:** Gribkova, E. D., Ibrahim, B. A. E., & Llano, D. A. (2018). A novel mutual information estimator to measure spike train correlations in a model thalamocortical network. *Journal of Neurophysiology*.

calcium current conductance influences mutual information between the input and output from this network. In particular, a T-type calcium current conductance of about 40 nS appears to produce maximal mutual information between the input to this network (conceptualized as afferent input to the thalamocortical cell) and the output of the network at the level of a layer 4 cortical neuron. Furthermore, at particular combinations of inputs to thalamocortical and thalamic reticular nucleus cells, thalamic cell bursting correlated strongly with recovery of mutual information between thalamic afferents and layer 4 neurons. These studies suggest that the novel mutual information estimator has advantages over previous estimators, and that thalamic reticular nucleus activity can enhance mutual information between thalamic afferents and thalamorecipient cells in the cortex.

INTRODUCTION:

Estimating the degree to which a spike train signal is conserved or transformed as it passes through a neural network requires methods that are both sensitive to spike timing and that take into account nonlinear dependencies between spike trains. These requirements are particularly necessary in thalamocortical networks, where cortical responses are not easily predicted from thalamic responses (Kayser et al., 2001; MacLean et al., 2005; Miller et al., 2014; Watson et al., 2008) and rely on the precise timing of inputs to process sensory information (Rose & Metherate, 2005). Sensory thalamocortical (TC) neurons receive input from retina or caudal sensory structures and project to well-defined areas of the cerebral cortex. TC neurons display at least two firing modes: tonic and burst mode, and tonic mode has generally been regarded as a high-fidelity transmission state to relay sensory information to the cortex (Jones, 2007; Kim & McCormick, 1998; Llinás & Steriade, 2006; D. McCormick & Feeser, 1990;

Sherman, 2001; Mircea Steriade & Llinás, 1988). TC cells are also strongly synaptically interconnected with the thalamic reticular nucleus (TRN), which comprises a shell of GABAergic neurons that partially surrounds the thalamus. The TRN has been implicated in a wide range of brain functions, including the production of sleep spindles, modulation of arousal and attention and, under pathological conditions, production of absence seizures (A. Destexhe et al., 1993; Alain Destexhe et al., 1999; Michael M. Halassa et al., 2014; M. M. Halassa et al., 2011; Huguenard, 1998; McAlonan et al., 2008; D. A. McCormick & Contreras, 2001).

Traditional models of thalamic processing have postulated the presence of reciprocal connectivity between TC and TRN neurons, and that this connectivity forms the basis for oscillations such as spindles and spike-wave discharges in absence epilepsy (Alain Destexhe et al., 1998; Huguenard, 1998; M Steriade et al., 1993). However, more recent data have revealed additional non-reciprocal connectivity between thalamocortical neurons and TRN, which may serve as a substrate to adjust the gain and filter properties of TC cells or to select particular groups of TC cells to meet cognitive demands (Crabtree & Isaac, 2002; Kimura, 2014; Pinault & Deschênes, 1998; Willis et al., 2015; Zikopoulos & Barbas, 2012). In our previous work (Willis et al., 2015), we explored a non-reciprocal ("open-loop") model and found that TRN stimulation did not universally depress TC spiking activity, but rather altered the cortical spike counts and mutual information between TC input and layer 4 (L4) cortical cell spiking in a frequency-dependent manner. Specifically, we found that total spike counts in the output of the model, an L4 cortical cell, were paradoxically enhanced by intermediate rates of TRN activity (approximately 25 Hz), and that this enhancement was dependent on both the actions of TRN neurons and nonlinear T-type calcium currents (T-currents) in TC cells, consistent with recent physiological findings (Whitmire et al., 2017; Whitmire et al., 2016). In contrast to our current

study, this previous work (Willis et al., 2015) provided only a limited exploration of the effect of variable TC T-current conductances on information flow in the open-loop thalamocortical network, especially since estimation of mutual information had not been employed between thalamic afferents and TC output of the thalamocortical network. Given the impact of T-channel-mediated bursting behavior on enhancement of information transmission through a TC network (Reinagel et al., 1999), and the theoretical benefits of bursting on signal encoding (Denning & Reinagel, 2005; Lisman, 1997; Mukherjee & Kaplan, 1995; Oswald et al., 2007; Person & Perkel, 2005; Reinagel et al., 1999; G. D. Smith et al., 2000; H. Swadlow & A. Gusev, 2001), we hypothesized that the TRN in an open-loop configuration could have a major impact on information traveling through the TC network.

There is, however, difficulty in measuring information flow through a network of spiking neurons, particularly in networks that must be able to respond to trains of incoming sensory signals at a broad range of rates, such as thalamocortical networks. There exist several generalized methods of estimating dependence between time series (see (Doquire & Verleysen, 2012; Silverman, 1986; Walters-Williams & Li, 2009) for overviews of common estimators). One such class of methods is distance metrics. The Victor-Purpura spike train distance metric (Victor & Purpura, 1996) is a binless method of measuring dissimilarity between spike trains. Unlike other distance metrics, this method embeds data in a metric space instead of a vector space, thus avoiding the assumption of spike train addition or scalar multiplication. Another common method is correlation analysis, which includes Pearson correlation and the Spike Timing Tiling Coefficient (STTC) (Cutts & Eglen, 2014). Of these methods, STTC was specifically developed for estimating correlation between neural spike trains, and it is not confounded by firing rate, unlike other measures of correlation.

However, a potential problem with traditional correlational analyses is that they are not optimal at estimating nonlinear dependencies (Gencaga et al., 2014), which are often observed in neural networks. This drawback can be addressed by using a dependence measure known as mutual information (MI) (Cellucci et al., 2005; Shannon, 1948). MI between two random variables X and Y is mathematically defined as:

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \quad , \tag{4.1}$$

where *H(X)* and *H(Y)* are entropies and *H(X,Y)* is joint entropy. Since the entropy *H* of a random variable $X = x_1, x_2, \ldots, x_n$ is essentially a measure of its unpredictability, and is defined through its probability distribution *p* with $H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i)$, MI estimators can take into account any kind of dependency. Furthermore, an important feature of MI estimators is that they can be used as nonparametric density estimators, as MI is considered to be a nonparametric measure of relevance (Walters-Williams & Li, 2009). Nonparametric estimators assume no prior model underlying the data distribution, and consequently require more data than parametric estimators (Worms & Touati, 2016). A common example of a nonparametric estimator is the Direct Method (Strong et al., 1998). This estimator calculates MI using the distribution of binary "words" in input and output time series. However, like many estimators, the Direct Method uses a uniform partition of the data, which consists of a fixed time window to obtain probability distributions. Uniform partitions are typically ad hoc or chosen through an error-reduction algorithm which increases computational cost (Kjaer et al., 1994; Walters-Williams & Li, 2009). By using adaptive partitions (Cellucci et al., 2005; Marek & Tichavsky, 2008), convergence of the MI estimate is faster, and the amount of data required is less, making adaptive partition methods generally more computationally efficient than uniform partitions, particularly when data

are distributed nonuniformly (Darbellay & Vajda, 1999; Marek & Tichavsky, 2008; Walters-Williams & Li, 2009).

In the case of thalamocortical networks, whose output depends heavily on the rate of synaptic input from both peripheral sensory structures and the TRN (Bartlett & Wang, 2007; Kim et al., 1997; Mukherjee & Kaplan, 1995; Willis et al., 2015), measurement of the rates of information flow across the thalamus require MI estimators that provide estimates across different spike rates. Therefore, we introduce an estimator of MI between time series (Adaptive partition using Interspike intervals MI Estimator, or AIMIE) that takes into account these different characteristic timescales. While AIMIE is similar to the estimator introduced in our previous work (Willis et al., 2015), the primary difference between these two estimators is the specific type of adaptive partition used and the way in which it is established. AIMIE, unlike previous estimators, uses adaptive partitions of interspike intervals and spikes densities to handle the disparate firing rates. Previously used estimators are typically limited by poor estimation of nonlinear dependencies, a priori assumptions of distributions, requirements of large amounts of data, and overestimation or underestimation of data distributions. In the current study, the AIMIE method is further examined and compared to these other, more traditional, methods of estimating MI using both simulated and real spike trains. AIMIE is then used to probe the impact of the TC T-current and the TRN on the transformation of spike information as signals pass through a simple thalamocortical network model. Using several different criteria, we find that AIMIE outperforms the other metrics and reveals that TRN-mediated inhibition in a thalamocortical model produces a paradoxical recovery of information per spike that is lost during thalamic bursting.

METHODS:

Computational Methods

Model Architecture:

A Hodgkin-Huxley framework was used to build a three-neuron, open-loop thalamoreticular network, as described previously (Willis et al., 2015). The network consists of single TC, TRN, and L4 cells, modeled as single-compartment models from whole-cell recordings done in our laboratory. Each cell's membrane potential V was modeled by a first-order differential equation:

$$C_m \frac{dV}{dt} = -g_L(V - E_L) - \sum_{i=1}^{n} g_i(V)(V - E_i) + I_e \,, \qquad (4.2)$$

where C_m is membrane capacitance, g_L is leak conductance, E_L is leak reversal potential, I_e is current externally applied to the cell, n is the number of channel types, excluding leak channels, $g_i(V)$ is conductance of ith channel type as a function of the membrane potential, and E_i is reversal potential of i^{th} channel type.

As described previously (Willis et al., 2015), the model of TC cell includes T-type calcium current, cationic H-current, delayed-rectifying potassium current, and fast sodium current, while the TRN cell model includes all of the aforementioned currents as well as slow-inactivating potassium current (KS current). To detect the maximum number of bursts, the number of bursts in TC output time series was calculated using liberal criteria, defined by (Ramcharan et al., 2000) as, for each burst, at least 50 ms of quiescence followed by at least two spikes with interspike interval(s) of at most 6 ms (but see (Deleuze et al., 2012; Sincich et al., 2007)). The L4 model includes fast sodium current, delayed-rectifying potassium current, and non-inactivating potassium current (M-current). The parameters of these cell models and their currents are found in **Table 4.1**. Thalamic afferent, reticulothalamic, and thalamocortical

synaptic parameters were derived from the literature (Chen & Regehr, 2003; Gentet & Ulrich, 2003; Laurent et al., 2002). All inputs to TC and TRN cell models are Poisson-modulated pulse trains, with a single-pulse duration of 0.1 ms.

Table 4.1. Thalamocortical Network Model Cellular Parameters

TC Cell	
Leak conductance	3.263 nS
Leak equilibrium potential	-60.03 mV
Fast sodium conductance	1500 nS
Sodium equilibrium potential	50 mV
Potassium conductance	520 nS
Potassium equilibrium potential	-100 mV
T-current conductance	45 nS
Calcium equilibrium potential	120 mV
H-current conductance	0.608 nS
H-current equilibrium potential	-33 mV
L4 Cell	
Leak conductance	4.8128 nS
Leak equilibrium potential	-60.2354 mV
Fast sodium conductance	3000 nS
Sodium equilibrium potential	50 mV
Potassium conductance	140 nS
Potassium equilibrium potential	-90 mV
M-current conductance	1.5 nS
TRN Cell	
Leak conductance	3.7928 nS
Leak equilibrium potential	-57 mV
Fast sodium conductance	3000 nS
Sodium equilibrium potential	50 mV
Potassium conductance	400 nS
Potassium equilibrium potential	-100 nS
T-current amplitude	21 nS
H-current amplitude	0.0192 nS
H-current equilibrium potential	-33 mV
KS-current conductance	3.5 nS
KS current τ	200 ms

MI Estimator with Adaptive Partition based on Interspike Intervals (AIMIE):

AIMIE is a nonparametric MI estimator (Walters-Williams & Li, 2009), which does not assume the distribution of data a priori, similar to the MI estimators mentioned below, and utilizes an adaptive partition (Cellucci et al., 2005) of interspike interval durations and spike densities. Given two time series of spike times, let X be the time series with the greater number of spikes, and Y be the time series with the lesser number of spikes. Let t_i^Y be the times at which

spikes occur in Y, and let Δt_i^Y be the duration (in ms) of an interval between subsequent spikes of Y such that

$$\Delta t_i^Y = t_{i+1}^Y - t_i^Y \quad . \tag{4.3}$$

Let n_i^X be the number of spikes of X in the Y interspike interval $[t_i^Y, t_{i+1}^Y)$ corresponding to Δt_i^Y (**Figure 4.1A**), and M be the total number of Y interspike intervals. From Y we constructed a time series of interspike interval durations, Δt_1^Y, Δt_2^Y, …, Δt_M^Y, and from X we constructed a time series of spike densities, d_1^X, d_2^X, …, d_M^X, corresponding to each interspike interval of Y, where $d_i^X = \frac{n_i^X}{\Delta t_i^Y}$.

An adaptive partition was applied to the time series of interspike interval durations of Y and separately to the time series of spike densities of X (**Figure 4.1B**). Specifically, for the adaptive partition, the corresponding time series was sorted by smallest to largest values, and each bin was chosen to have a minimum occupancy C_{min} equal to the square root of the total number of interspike intervals of Y, with the total number of bins, for either X or Y, being N. Let C_j^Y be the occupancy of bin j of the adaptive partition used for Y, and let C_k^X be the occupancy of bin k of the adaptive partition used for X. For each bin, the marginal probabilities, P_j^Y and P_k^X, were calculated as follows (**Figure 4.1B**):

$$P_j^Y = \frac{C_j^Y}{\sum_{l=1}^{N} C_l^Y} \ , \tag{4.4}$$

$$P_k^X = \frac{C_k^X}{\sum_{l=1}^{N} C_l^X} \ . \tag{4.5}$$

Note that the denominator for each marginal probability is the sum of all occupancies for the corresponding adaptive partition. This particular partition, where each bin is nonempty and has roughly equal occupancy, is related to an adaptive partition (Cellucci et al., 2005).

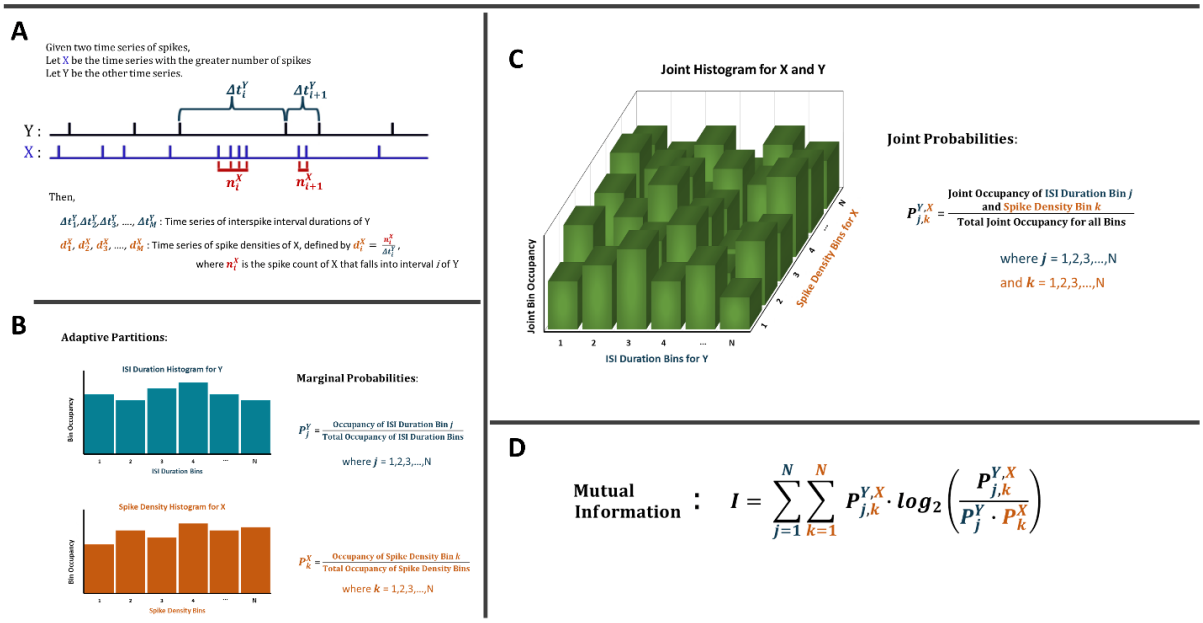

Figure 4.1. Procedure for estimating MI using AIMIE. **A)** Given two time series, the one that has a greater number of spikes is designated as X, and the other is designated as Y. A time series of interspike interval durations for Y is constructed, and a time series of spike densities of X corresponding to the interspike intervals of Y. M represents the total number of interspike intervals of Y. **B)** An adaptive partition is applied to the time series of interspike interval durations of Y and separately to the time series of spike densities of X. Marginal probability for each bin of the adaptive partition is calculated as the occupancy of the bin divided by the sum of occupancies of all bins of that adaptive partition. Note the roughly equal occupancy of the bins, which is due to the adaptive partition. **C)** A joint histogram is constructed, in which one of the horizontal axes represents the bins of the adaptive partition for Y and the other horizontal axis represents bins of the adaptive partition for X. Joint probability for each combination of bins is calculated as joint occupancy of both bins divided by the sum of all occupancies of the joint histogram. **D)** Equation for calculating MI, in which the outer sum, from j=1 to N, sums over the bins of the adaptive partition of Y and the inner sum, from k=1 to N, sums over the bins of the adaptive partition of X.

The joint probability $P_{j,k}^{Y,X}$ was calculated by constructing a two-dimensional matrix with one horizontal axis corresponding to bins of the adaptive partition for Y and the other horizontal

axis corresponding to bins of the adaptive partition for X (**Figure 4.1C**). Let $C_{j,k}^{Y,X}$ be the joint occupancy of both bin j of the adaptive partition used for Y and of bin k of the adaptive partition used for X. Then $P_{j,k}^{Y,X}$ was calculated as follows:

$$P_{k,j}^{Y,X} = \frac{C_{k,j}^{Y,X}}{\sum_{l=1}^{N}\sum_{p=1}^{N} C_{l,p}^{Y,X}}, \qquad j = 1,2,3,...,N \text{ and } k = 1,2,3,...,N. \tag{4.6}$$

To calculate mutual information, the general formula for MI is used (**Figure 4.1D**):

$$I = \sum_{j=1}^{A}\sum_{k=1}^{B} P_{j,k}^{Y,X} \log_2 \frac{P_{j,k}^{Y,X}}{P_j^Y P_k^X} \qquad j = 1,2,3,...,N \text{ and } k = 1,2,3,...,N, \tag{4.7}$$

where N, being the total number of bins for either adaptive partition in this case also served as the total number of bins on both axes, such that A = B = N.

Direct Method MI estimator (DMIE):

This estimator uses a fixed-width partition of binary input and output time series (Strong et al., 1998). This technique has a precision of Δt, which is the sampling interval that is used to convert input and output time series into binary signals. A time window of length T was used to construct a dictionary of unique binary "words," each consisting of binary symbols (0's and 1's), from the input and output. Marginal probabilities, P_i^{out} and P_j^{in}, associated with output word i and input word j, were calculated by dividing the total number of occurrences for each word by the total number of occurrences for the corresponding time series. Joint probability, $P_{i,j}^{out,in}$, was determined by constructing a two-dimensional matrix of output words, corresponding to columns, and input words, corresponding to rows, much like with AIMIE. The occupancy of each two-dimensional bin was then divided by the total occupancy of all two-dimensional bins to give joint probability for the corresponding two-dimensional bin. Finally, MI was calculated

using **Eq. (4.7)**, with B as the total number of output words, and A as the total number of input words. We utilized two variations of DMIE: DMIE10, with $\Delta t = 10$ ms and $T = 100$ ms, and DMIE20, with $\Delta t = 20$ ms and $T = 200$ ms.

These partition parameters were chosen such that in either variation, the binary words would consist of 10 symbols, as recommended by Strong et al., and to fully test the Direct Method's dependence on interspike interval length variation.

MI estimator with uniform partition of spike counts (three variations: FBWSE, FBNSE, SQRSE):

This is a simple estimator that uses a fixed bin-width partition to determine spike count distributions for input and output. Originally, this method of estimation was used to measure the amount of information transmitted by neuronal responses in the visual system about a set of stimuli (M. Oram et al., 1999; M. W. Oram et al., 2001). Specifically, this estimator can be used to compare information transmitted by total spike count as well as by number of repeating triplets in output spike code. For this method, a uniform partition with time bins of fixed duration, also used for DMIE10, was used to construct a time series of spike counts. Marginal probabilities, P_i^{out} and P_j^{in}, were determined from histograms that were constructed based on spike counts, with spike bin occupancy defined as the number of time bins containing the corresponding number of spikes. Joint probability, $P_{i,j}^{out,in}$, was determined by constructing a two dimensional histogram of input spike bins on one axis and output spike bins on the other axis. MI was calculated using **Eq. (4.7)**, with B as the total number of output spike bins, and A as the total number of input spike bins.

To determine distribution histograms, Oram et. al used the error reduction algorithm of Kjaer et. al (Kjaer et al., 1994; M. Oram et al., 1999; M. W. Oram et al., 2001), which requires training a back-propagation artificial neural network with relatively substantial amounts of data. However, to demonstrate a blind application of this estimator, we defined three variations of uniform partitions that differ only in their construction of time bins. The fixed bin-width spike partition (FBWSE) uses time bins with a fixed duration of 10 ms, regardless of the duration of input or output time series. The fixed bin-number spike partition (FBNSE) uses 500 time bins for both input and output, with bin width being the total duration of the time series divided by number of bins. For the square-root spike partition (SQRSE), we chose the number of time bins to be the square root of the total number of data points for each time series (Cellucci et al., 2005; Mosteller & Tukey, 1977). The square-root partition may provide an optimal bin-width for constructed histograms, as suggested by Mosteller and Tukey.

All simulations and estimators were coded in MATLAB R2012a and R2015a. MI was computed using the algorithms of Celluci et. al (Cellucci et al., 2005), related to the plug-in method of MI estimation (Paninski, 2003). Calculations and simulations were run on multiple machines, including an HP Pavilion machine using a Windows 7 operating system, as well as a Lenovo Ideapad machine with a Windows 10 operating system.

The MATLAB code for the thalamocortical model and AIMIE has been made publicly available. This code can be accessed via GitHub at https://github.com/KatyaGribkova/TCmodel for the thalamocortical model and at https://github.com/KatyaGribkova/AIMIE for AIMIE.

Calculation of Exact MI between Random Variables:

Exact MI was calculated between several pairs of random variables. Let X_i and N be normally distributed random variables with means (in ms) $\mu_i = 5i$ and $\mu_N = 5$, and standard deviation (in ms) $\sigma_i = i$ and $\sigma_N = 1$, respectively, with i = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Let Y_i be the sum of two random variables such that $Y_i = X_i + N$ for i = 1, 2, …, 10. Then the exact MI between X_i and Y_i is calculated as $I(X_i, Y_i) = \frac{1}{2}\log_2\left(1 + \frac{\sigma_i^2}{\sigma_N^2}\right)$. A pair of spike time series, consisting of 1000 spikes for each series, was then generated from the random variables X_i and Y_i using MATLAB's normrnd function, and the MI between them was estimated using the MI estimators AIMIE, DMIE10, DMIE20, FBNSE, FBWSE, and SQRSE. 10 trials were performed for spike time series pairs generated from X_i and Y_i, for for i = 1, 2, …, 10. For each estimator, the root mean square error (RMSE) was computed between the MI estimate and exact MI for each random variable pair, and an average RMSE was calculated over all random variable pairs X_i and Y_i for i = 1, 2, …, 10.

Statistical Analysis and Curve Fitting:

For datasets, normality was tested using the Shapiro-Wilk test. In all cases, the Shapiro-Wilk test was significant, indicating deviation from normality, and therefore Kruskal-Wallis and post-hoc Nemenyi tests were utilized. For simulation time datasets, standard deviation data points were fitted to an exponential curve for each MI estimator. MI datasets obtained from simulations of the thalamocortical model across different T-currents were fitted to logistic curves of the form $f(x) = \frac{a}{1 + e^{b\cdot(x-c)}} + d$ via the MATLAB function lsqcurvefit where a = the height of the curve, or the difference between maximum and minimum curve values, b = steepness of

the logistic curve, c = the T-current conductance midpoint of the curve, and d = the minimum value of the logistic curve. All fitted curves were tested for goodness of fit using the standard error of estimate metric, with each fit providing at least a 15% decrease in standard error of estimate when tested with a full dataset vs. half of the dataset. All curve-fitting and statistical analyses were performed in MATLAB 2016b and Excel 2016 with Real Statistics Resource Pack.

Electrophysiological Recordings

Animals:

P20-24 BALB/c mice of both sexes were used for this study. All procedures were approved by the Institutional Animal Care and Use Committee (IACUC, protocol # 16164) at University of Illinois Urbana-Champaign. Animals were housed in animal care facilities at the Beckman Institute for Advanced Science and Technology, approved by the American Association for Accreditation of Laboratory Animal Care (AAALAC).

Brain slicing:

Mice were initially anesthetized with ketamine (100 mg/kg) and xylazine (3 mg/kg) intraperitoneally and perfused with chilled (4^{o}C) sucrose-based slicing solution ((in mM): 234 sucrose, 11 glucose, 26 $NaHCO_3$, 2.5 KCl, 1.25 NaH_2PO_4, 10 $MgCl_2$, 0.5 $CaCl_2$). 300 μm-thick thalamocortical brain slices were obtained, as previously described (Cruikshank et al., 2002; Stebbings et al., 2016), then incubated for 30 minutes in 32^{o}C incubation solution ((in mM): 26 $NaHCO_3$, 2.5 KCl, 10 glucose, 126 NaCl, 1.25 NaH_2PO_4, 3 $MgCl_2$, and 1 $CaCl_2$). After incubation, slices were transferred to a perfusion chamber and perfused with artificial

cerebrospinal fluid (ACSF) ((in mM): 26 $NaHCO_3$, 2.5 KCl, 10 glucose, 126 NaCl, 1.25 NaH_2PO_4, 2 $MgCl_2$, and 2 $CaCl_2$), and bubbled with 95% oxygen/5% carbon dioxide.

Electrophysiology:

The cell-attached recordings of pairs of neurons located in layer 2/3 or layer 4 of the auditory cortex were performed at room temperature using a visualized slice setup outfitted with infrared-differential interference contrast (IR-DIC) optics. Recording pipettes were pulled from borosilicate glass capillary tubes and had tip resistances of 2–5 MΩ when filled with solution, which contained (in mM): 117 K-gluconate, 13 KCl, 1.0 $MgCl_2$, 0.07 $CaCl_2$, 0.1 ethyleneglycol-bis (2-aminoethylether)-N,N,N0 ,N0 -tetra acetic acid, 10.0 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, 2.0 Na-ATP, 0.4 Na-GTP, pH 7.3. For data acquisition, a Multiclamp 700B amplifier (Molecular Devices, Sunnyvale, CA, USA) and pClamp software (Molecular Devices, Sunnyvale, CA, USA) were used with a 20-kHz sampling rate. The cell-attached recordings were conducted under after a gigaOhm seal was attained. Once the recording was started, 2 μM of SR-95531 (gabazine, Tocris) was perfused with the ACSF for 30 minutes, after which increasing concentrations of DNQX (Tocris) were added sequentially to the ACSF (approximately 30 minutes/each DNQX concentration) along with 2 μM SR-95531. The software program Clampfit (Molecular Devices, LLC) was used to analyze series of spike times from paired recording data using an event detection algorithm with threshold search.

RESULTS:

The validity of using AIMIE to measure the degree of correlation between spike trains was examined. Four other MI estimators, DMIE, FBWSE, FBNSE, and SQRSE, described in further detail in Methods, are also tested for comparison.

Comparison of MI estimates with exact MI values

Exact MI between two random variables can be calculated when their probability density functions are precisely known, along with their joint probability distribution. Spike time series provide only a limited sample of the random variables that generate them, therefore exact MI between spike time series cannot be calculated, but can be estimated using MI estimators such as AIMIE. To compare AIMIE's estimate of MI with exact MI, spike time series were generated from sets of distributions with exactly calculable MI. Specifically, a series of random variable pairs was used, with exact MI calculated for each pair as detailed in Methods, as well as spike time series consisting of 1000 spikes generated from each random variable pair for MI estimators. For each MI estimator, AIMIE, DMIE10, DMIE20, FBNSE, FBWSE, and SQRSE, the average RMSE was computed between the MI estimates and exact MI (**Figure 4.2**). **Figure 4.2** shows that AIMIE has the lowest average RMSE, followed by DMIE20 and DMIE10 (Kruskal-Wallis and Nemenyi tests, $p = 0.0195$ for AIMIE vs. DMIE20). This finding indicates that AIMIE provides MI estimates that are, on average, closest to the exact MI among the tested estimators. Note that when the RMSE for AIMIE is compared across the random variable pairs, AIMIE overestimated MI for several random variable pairs which have smaller means and standard deviations, and underestimated MI for random variable pairs that have higher means and standard deviations. This indicates that AIMIE shows either positive or negative bias, depending on the data and underlying distributions used.

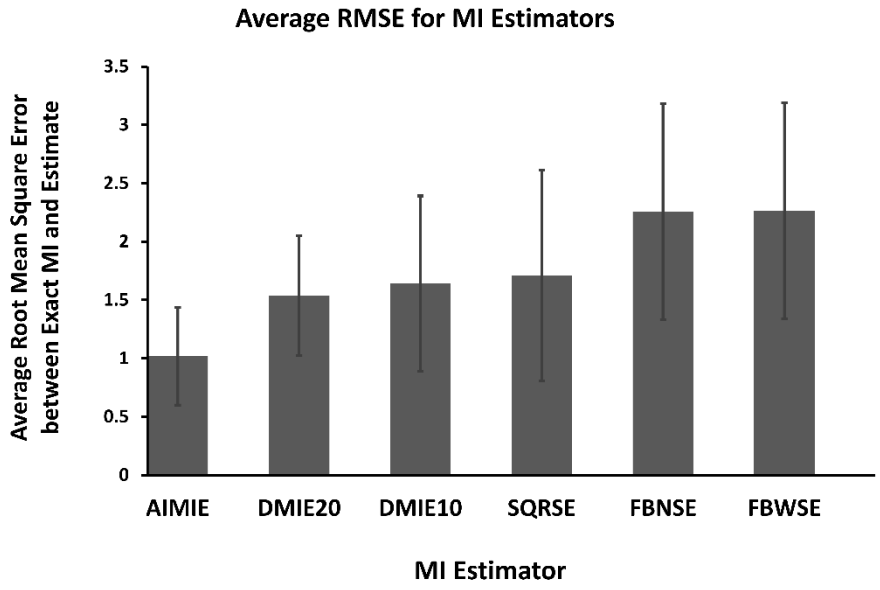

Figure 4.2. Comparison of RMSE of MI Estimators. MI was estimated between spike time series generated from random variables X_i and Y_i (i = 1, 2, …, 10), and this estimate was compared to the exact MI between X_i and Y_i. For each i = 1, 2, …, 10, the random variable Y_i is a summation of X_i, a normally distributed random variable with mean and standard deviation of 5i and i (in ms), respectively, and N, another normally distributed random variable which serves as Gaussian noise with a mean of 5 ms and standard deviation of 1 ms. 10 trials were used for generating spike time series for MI estimation. For each MI estimator, the RMSE was calculated between each MI estimate and exact MI value and then averaged over all random variable pairs. The average RMSE was compared across all MI estimators.

MI estimator dependence on the length of time series

To determine the number of spikes needed to achieve a stable estimate of MI using AIMIE, simulations of a thalamocortical model were run while systematically varying the simulation time. AIMIE was used to compute the MI between the afferent spike train providing synaptic input to a model TC neuron and the output of the model, measured as spike times in a model layer 4 cortical neuron.

A simple thalamocortical network model was used (**Figure 4.3A**, see **Table 4.1** for parameters), and the afferent input was a 10 Hz Poisson-modulated spike train, varying only in the total simulation time of the model. This scheme permitted generation of pairs of input and

output time series to examine basic properties of AIMIE, such as the dependency of AIMIE on the simulation time, which corresponds to total number of output spikes. Ten trials were performed for each simulation time to calculate standard deviation. The large error bar for the MI per output spike is due to the small numbers of spikes generated for short simulation times. For instance, for simulation times of 1 second, the number of output spikes ranges from 2 to 5. In the case where there are 2 or less output spikes, AIMIE estimated MI values to be zero. Comparing across the total simulation times, MI per output spike approaches an asymptotic value of about $2x10^{-4}$ bits/spike (**Figure 4.3B**). This finding indicates that with increasing amounts of data, AIMIE becomes less dependent on time series length. At about 100 seconds of simulation time (black triangle of **Figure 4.3**), which corresponds to an average of 411 output spikes, a decrease is seen in the standard deviation function (**Figure 4.3C**) and in magnitude of error bars (**Figure 4.3B** inset) when compared to shorter simulation times. Note that for simulation times below 100 seconds there are significant fluctuations in the MI values, however the standard deviation substantially drops at simulation times beyond 100 seconds. Therefore, for most subsequent simulations, simulation times were adjusted such that at least 500 output spikes were generated, and for all tests of AIMIE's performance, with artificial spike trains that were not generated with the thalamocortical model, a minimum of 500 output spikes was also used.

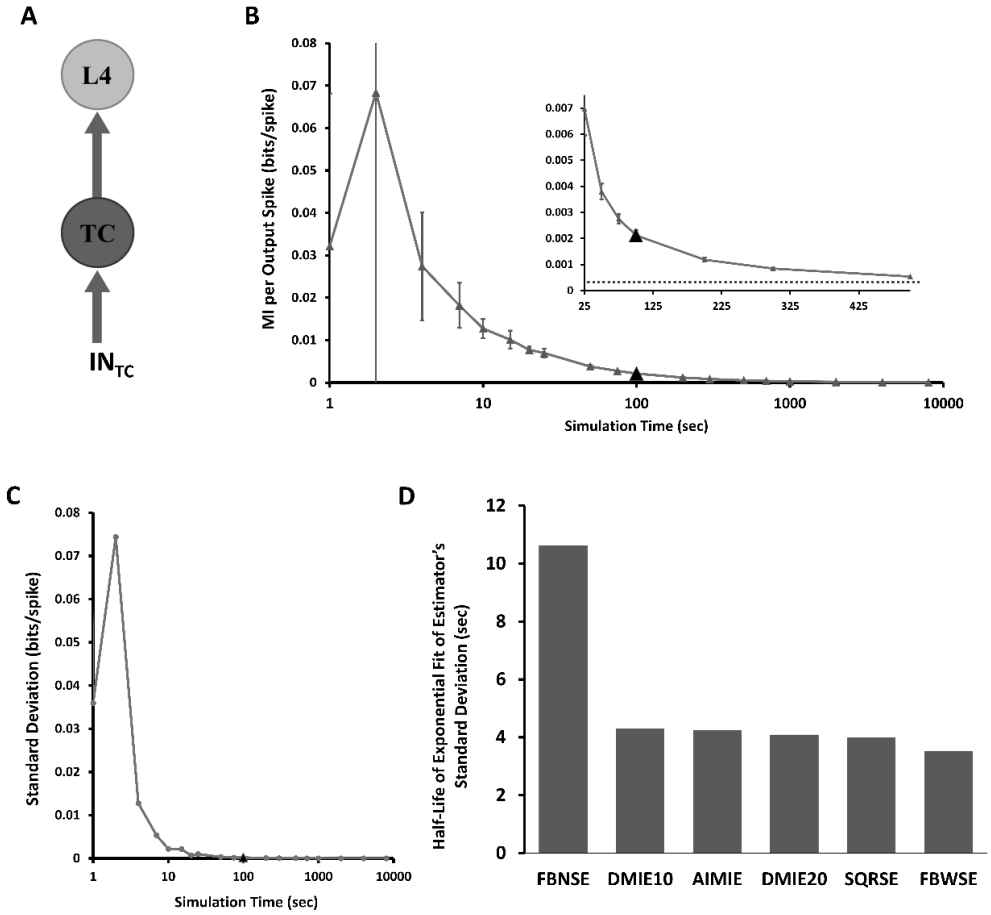

Figure 4.3. Application of AIMIE and other MI estimators to time series generated from the model thalamocortical network with variable simulation times. **A)** Model architecture of a thalamocortical network containing only TC and L4 neurons, which are modeled using a Hodgkin-Huxley framework. Default model parameters are used for simulations (see **Table 4.1**). INTC corresponds to thalamic afferent inputs, which are generated as 10 Hz Poisson-modulated pulse trains. **B)** MI per output spike provided by AIMIE when it is applied to time series pairs with different simulation times. As simulation time increases, the number of output spikes increases as well, and AIMIE's MI per output spike trends towards a horizontal asymptote of about 2x10-4 bits/spike. The black triangle marks a simulation time of 100 seconds, which provides an average of about 411 output spikes. Inset shows decreased fluctuations in MI per output spike on an expanded scale, particularly for simulation times above 100 seconds. 10 trials were used for each simulation time, to generate error bars and standard deviation. **C)** Standard deviation of AIMIE's estimates of MI per output spike from Figure 4.3A for a range of simulation times. Note that standard deviation is significantly smaller at around 100 seconds of simulation time (black triangle) than at shorter simulation times. For all other estimators, MI per output spike for the same set of simulation times was calculated, as well as the standard deviation of their MI estimates, just as for Figure 4.3B-C. For each MI estimator, the standard deviation of the MI estimate across simulation times 4 secs to 2000 secs was fitted to an exponential curve, followed by calculation of the simulation time at which the curve dropped to **D)** one-half of its initial value.

The dependency of each MI estimator on the simulation time was examined using the same set of simulated spike trains as with AIMIE, and the standard deviations of MI values for simulation times ranging from 4 seconds to 2000 seconds were fitted to an exponential curve. The simulation times at which the fitted exponential curve dropped to one-half of the exponential curve's value at 4 seconds were calculated to compare the rate at which the standard deviation of MI values fell for each estimator (**Figure 4.3D**). This comparison of the relative rates at which the standard deviation falls across estimators shows that the amount of data required for AIMIE's estimate of MI is comparable to that of DMIE10 and DMIE20, and is less than that of FBNSE. This also suggests that as the amount of data used increases, AIMIE's bias is reduced more quickly than FBNSE's. SQRSE and FBWSE have the highest rate of convergence, and as such appear to require the smallest amount of data for their MI estimate.

MI estimator independence of interspike interval length variation

To determine whether any of the MI estimators are sensitive to the average spiking rate, we created artificial input and output spike trains and computed MI between these trains using multiple different MI estimators. The input time series were randomly generated using a uniform distribution, each with 4000 spikes in total and interspike intervals ranging from 5 ms to 50 ms. In this case, for the output time series, a single output spike followed each input spike and arrived before the next input spike. Under the assumption here that the optimal estimator for MI should be insensitive to stretch (and therefore average spiking rate), the input and output spike trains were then temporally "stretched" to change their average spiking rate without changing the relationship between the two trains (**Figure 4.4A**). In total, ten trials were performed for each average spiking rate. When tested alongside the four other estimators (**Figure 4.4B**), the estimators that produced a change in MI beyond 10% of the baseline MI were DMIE10,

DMIE20, and FBWIE, all of which use a fixed bin width partition, indicating that DMIE and FBWIE are sensitive to time scaling, while FBNIE, SQRSE, and AIMIE are not. Therefore, FBNIE, SQRSE and AIMIE were compared in subsequent examinations of MI estimator performance.

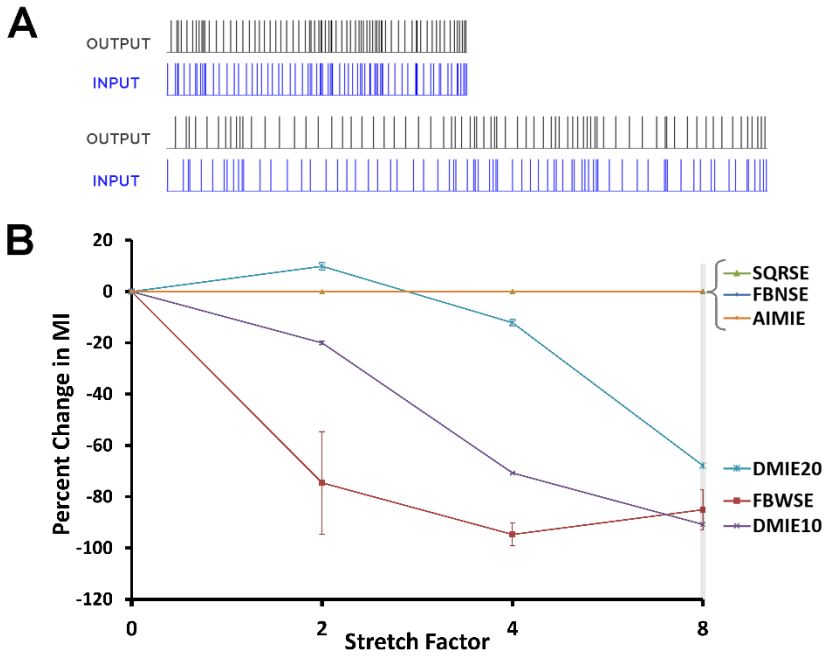

Figure 4.4. Effect of scaling of interspike intervals on MI for different estimators. **A)** Spike time plots demonstrating scaling of interspike intervals for input and output series. In this case, the second input and output pair is generated by scaling all interspike intervals of the first pair by a factor of 2. **B)** Percent change in MI demonstrated by estimators AIMIE, DMIE, FBWSE, FBNSE, and SQRSE when they are applied to input and output time series with interspike intervals scaled by factors of 2, 4, and 8. 10 trials were used for each data point and to generate standard deviation for error bars.

MI estimator performance with artificial spike trains

In the next test of MI estimator performance, MI estimators that performed well in the previous test were applied to time series pairs that were artificially generated to have varying degrees of dependence between input and output. This assessment allowed evaluation of whether the MI estimator could provide MI values that reflect the relative degree of dependence between

time series pairs. It was assumed that a high dependence between input and output data would be indicated by each input spike eliciting a similar spiking event in the output. Deviation from this idealized response would be considered a loss of information.

To test the performance of the MI estimators, four sets of input and output time series were generated, with progressively decreasing degree of correlation in the first three sets (**Figure 4.5A**). The input time series were randomly generated, each with 4000 spikes in total and interspike intervals ranging from 5 ms to 50 ms using a uniform distribution. For the first set of input and output time series, designated as Type 1, the response is ideal, meaning that each input spike has 100% chance of eliciting a single output spike. For the Type 2 pair, there is a 50% response, where each input spike has a 50% chance of generating a single output spike. This adjustment roughly decreases the number of output spikes in Type 2 by half when compared to Type 1, and as such, should constitute a decrease in MI. Similarly, the Type 3 pair features a 25% response. The pair of time series in Type 4 is constructed similarly to Type 3 time series, with the 4000 inputs randomly generated; each input has a 25% chance of generating a response event in the output, which, unlike for Type 3, is a random number between 1 to 10 spikes. Ten trials were performed for each application of an MI estimator to a different type of time series pair. Because Type 1 through Type 3 pairs show a progressive loss of spikes in the output time series and same number of spikes for the input time series, our expectation was that an ideal MI estimator would show the highest MI for Type 1 stimuli and progressively lower for Type 2 and Type 3 stimuli. If each burst in Type 4 output time series is considered as a single event, then it becomes comparable to Type 3 output time series, so there is also the expectation that an ideal MI estimator should show lower MI for both Type 3 and Type 4 stimuli than for Type 2 stimuli.

For this comparison, only the MI estimators that were shown to be insensitive to time scaling (FBNIE, SQRSE, and AIMIE) were tested. When AIMIE was applied to these four pairs of constructed time series, it was observed that the MI values corresponded to the drops in degree of correlation across Types 1 through 3 (**Figure 4.5B**).

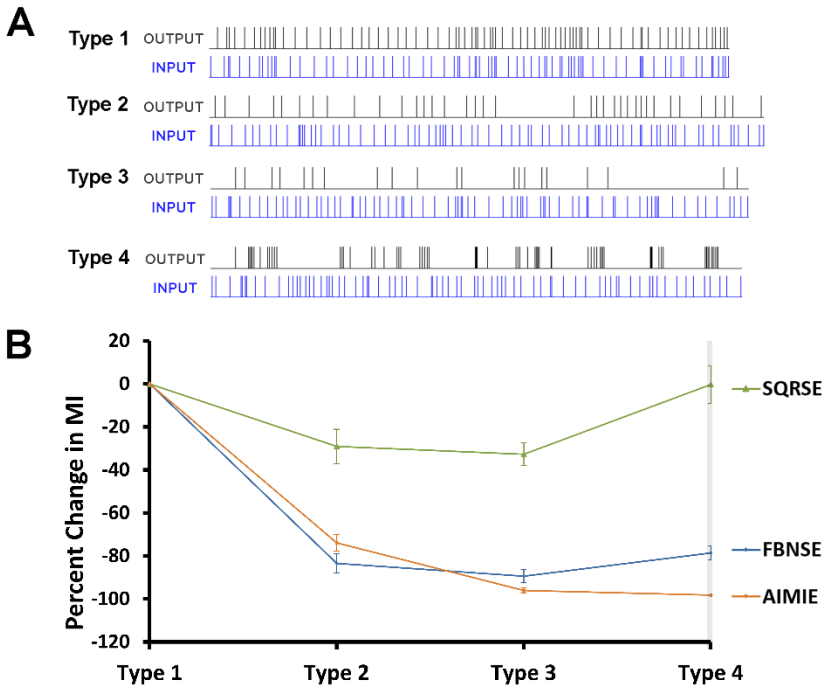

Figure 4.5. Application of different estimators to hypothetical input and output series of variable dependence. **A)** Spike time plots of hypothetical constructed time series of Types 1, 2 ,3, and 4, where Type 1 is an ideal (100%) response type, Type 2 is a 50% response type, Type 3 is a 25% response type, and Type 4 is a 25% burst response type. **B)** Percent change in MI demonstrated by estimators AIMIE, FBNSE, and SQRSE when they are applied to input and output time series of Types 1, 2, 3, and 4. As before, 10 trials were used for each data point and to generate standard deviation for error bars.

For Type 4, the MI value was very close to that of Type 3, within 2%. FBNSE and SQRSE, neither of which uses an adaptive partition, demonstrated a significant drop in MI across Type 1 to Type 2 and a slight drop in MI from Type 2 to Type 3. Lastly, for Type 4, when compared to Type 3, both FBNSE and SQRSE showed an increase in MI. Note that the only estimator whose MI value for Type 4 was less than that of Type 3 is AIMIE. The drop from Type 1 to Type 4 was

significantly different (Kruskal-Wallis and Nemenyi tests, $p = 0.0215$ for AIMIE vs. FBNSE and $p < 0.001$ for AIMIE vs. SQRSE) for AIMIE when compared to that of FBNSE and SQRSE. Substantial fluctuation in MI values, resulting in significant error values, may indicate an insufficient amount of data, though as suggested by our time series length test (**Figure 4.3**), this amount of data would be enough for AIMIE to yield reliable results, as there are 4000 input spikes and around 1000 or more output spikes and for each time series pair type. Note that the degree of variance for AIMIE's estimates is noticeably smaller than those of the other estimators' values. These data suggest that AIMIE is more sensitive than other methods to manipulations of spike trains that are expected to diminish the MI between them.

MI estimator performance with paired electrophysiological recordings

In this test, MI estimators that performed well in the previous analysis were applied to time series pairs with varying degrees of synchrony, which were obtained in electrophysiological experiments. In contrast to the previous test, these time series pairs showed greater variability in number and timing of spikes. The MI estimators were applied to spike trains of paired recordings of spontaneous activity from the auditory cortex in a slice preparation, under different concentrations of bath-applied DNQX, which is an AMPA receptor antagonist that blocks excitatory synaptic transmission. At higher concentrations of DNQX, we expect that MI per spike between responses of the two neurons of a paired recording would drop, due to diminished synchrony between spontaneous action potentials of the paired neurons. Specifically, AIMIE and FBNSE were tested, as these two estimators performed well with artificially constructed inputs and outputs described above. Recordings from a total of five pairs of neurons in layer 2/3 or layer 4 from mouse brain slices containing the auditory cortex were used.

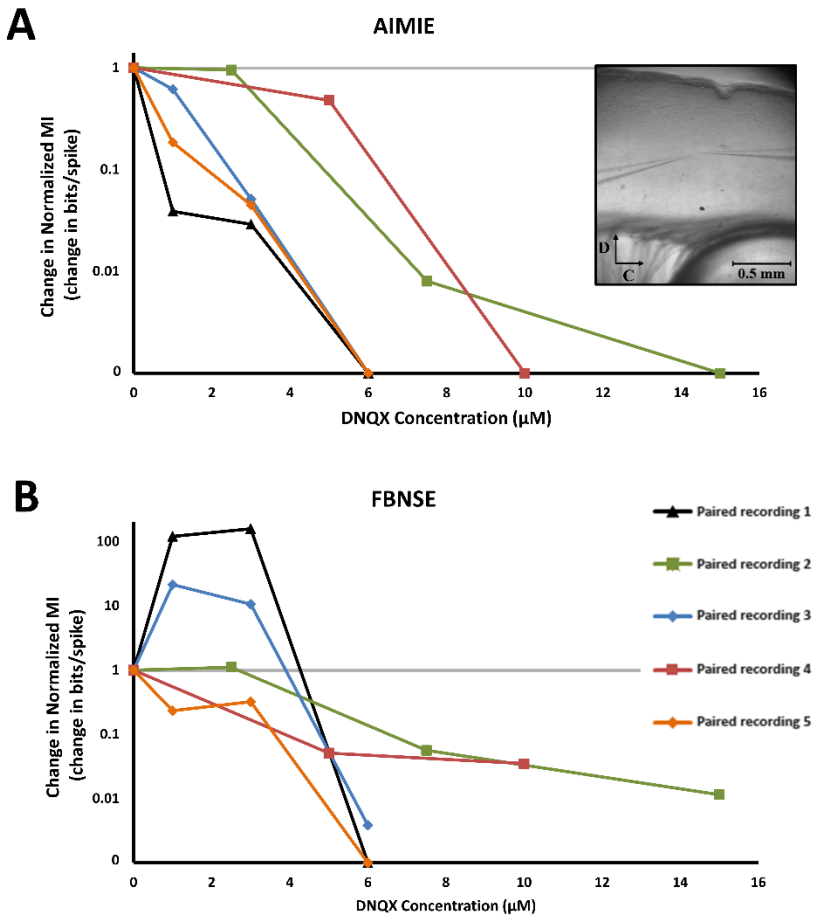

Figure 4.6. Application of different estimators to dual-recorded spike time series of variable synchrony. Recordings from a total of five neuron pairs are used. For both estimators, for each paired recording sequence under different DNQX concentrations, the change in normalized MI is calculated as MI per spike at each DNQX concentration divided by the MI per spike at DNQX concentration of 0 µM. **A)** Change in normalized MI demonstrated by AIMIE across different concentrations of DNQX. The inset is an image of electrode placement for a paired recording in auditory cortex of mouse brain slice, with C denoting caudal and D denoting dorsal orientation. The legend for paired recordings is located below in Figure 4.6B. **B)** Change in normalized MI demonstrated by FBNSE across different concentrations of DNQX. The legend corresponds to both Figure 4.6A and Figure 4.6B.

For each pair of recorded neurons, the neuron that produced the greatest number of spikes at the highest DNQX concentration was used for normalization of MI, specifically by dividing the MI value of each trial by the number of spikes produced by this neuron for that trial. For each paired recording sequence under different DNQX concentrations, the change in normalized MI

was calculated by dividing the MI per spike at each DNQX concentration by the MI per spike at DNQX concentration of 0 μM. With increasing DNQX concentrations, AIMIE demonstrated the expected drop in normalized MI (**Figure 4.6A**), while FBNSE did not (**Figure 4.6B**), indicating that AIMIE performs well with electrophysiological recordings of variable synchrony.

Varying T-current conductance in open-loop thalamoreticular network

Of the MI estimators tested, it was expected that AIMIE would be the most suitable estimator to apply to data generated from the thalamocortical model, since AIMIE was shown to have the best performance across the previous tests, and because data used in these previous tests shares several qualities with data generated from the model, such as variable degrees of dependence. Thus, applying AIMIE to a simple open-loop thalamocortical network model (**Figure 4.7A**), the effect of varying stimulation parameters and the T-current conductance of the TC cell on the modification of an ascending signal through the thalamus was explored.

Poisson-modulated synaptic input to a model TC cell (henceforth "thalamic afferents") was varied from 0.5 Hz to 200 Hz on a logarithmic scale, while Poisson-modulated synaptic input to a TRN neuron was similarly varied from 0 Hz to 200 Hz. As before (Willis et al. 2015), the applied spike trains in thalamic afferents were independent of the applied spike trains in afferents to the TRN cell. For each combination of TC and TRN stimulation rates, the MI between the thalamic afferents and output, either at the TC or L4 neuron, was computed. This MI, normalized by the number of output spikes, was then averaged over all combinations of TC and TRN stimulation rates to produce average MI per output spike. The average MI per output spike was computed for each T-current conductance, ranging from 0 to 100 nS (**Figure 4.7B**).

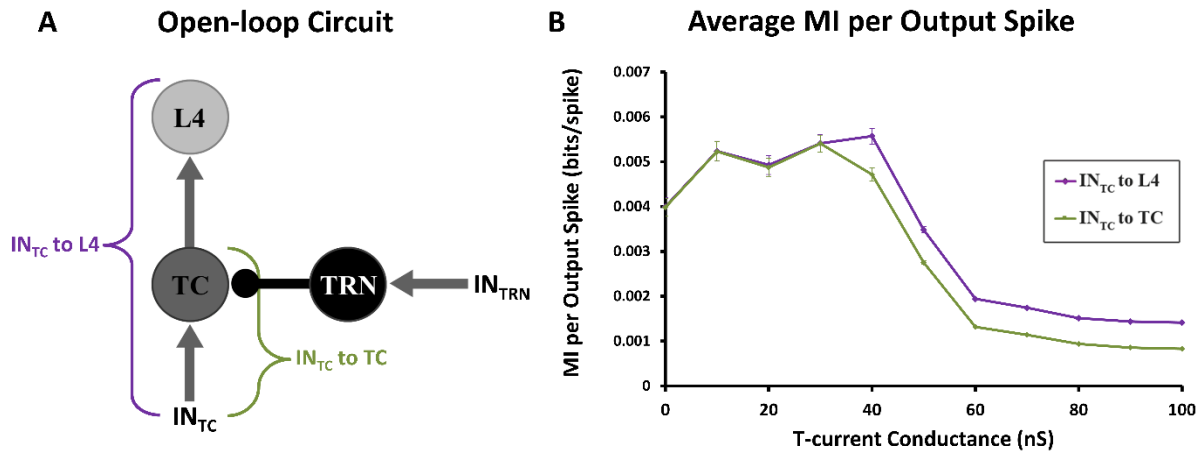

Figure 4.7. MI analysis of inputs and outputs in an open-loop thalamocortical network model at variable T-current conductances. **A)** Model architecture of the open-loop thalamocortical network. Arrows represent excitatory inputs, while the TRN to TC projection represents inhibitory (GABAergic) input. All neurons are modeled using a Hodgkin-Huxley framework. INTRN corresponds to input to the TRN, which ranges from 0 Hz to 200 Hz, while INTC corresponds to thalamic afferents, which ranges from 0.5 Hz to 200 Hz. Both INTRN and INTC are generated as Poisson-modulated pulse trains. Note that the green bracket symbolizes information transfer from thalamic afferents to TC, and the purple bracket symbolizes information transfer from thalamic afferents to L4. **B)** Average MI transmitted per output spike between thalamic afferents and TC and between thalamic afferents and L4, at a range of TC T-current conductances. There is a peak in MI per output spike at a TC T-current conductance of about 40 nS, which may indicate maximum potentiation of ascending input in open-loop thalamocortical network. 10 trials were used for each data point and generation of standard deviation for error bars. The average MI per output spike between the thalamic afferents and L4 is on average slightly greater than between the thalamic afferents and the TC cell. To examine the significance of this difference, each trial, consisting of MI per output spike values over all T-current conductances, was fitted to a logistic function of the form

$$f(x) = \frac{a}{1 + e^{b \cdot (x - c)}} + d$$

where a = the difference between the curve's maximum and minimum values, b = steepness of the logistic curve, c = the T-current conductance midpoint of the curve, and d = the minimum value of the logistic curve.

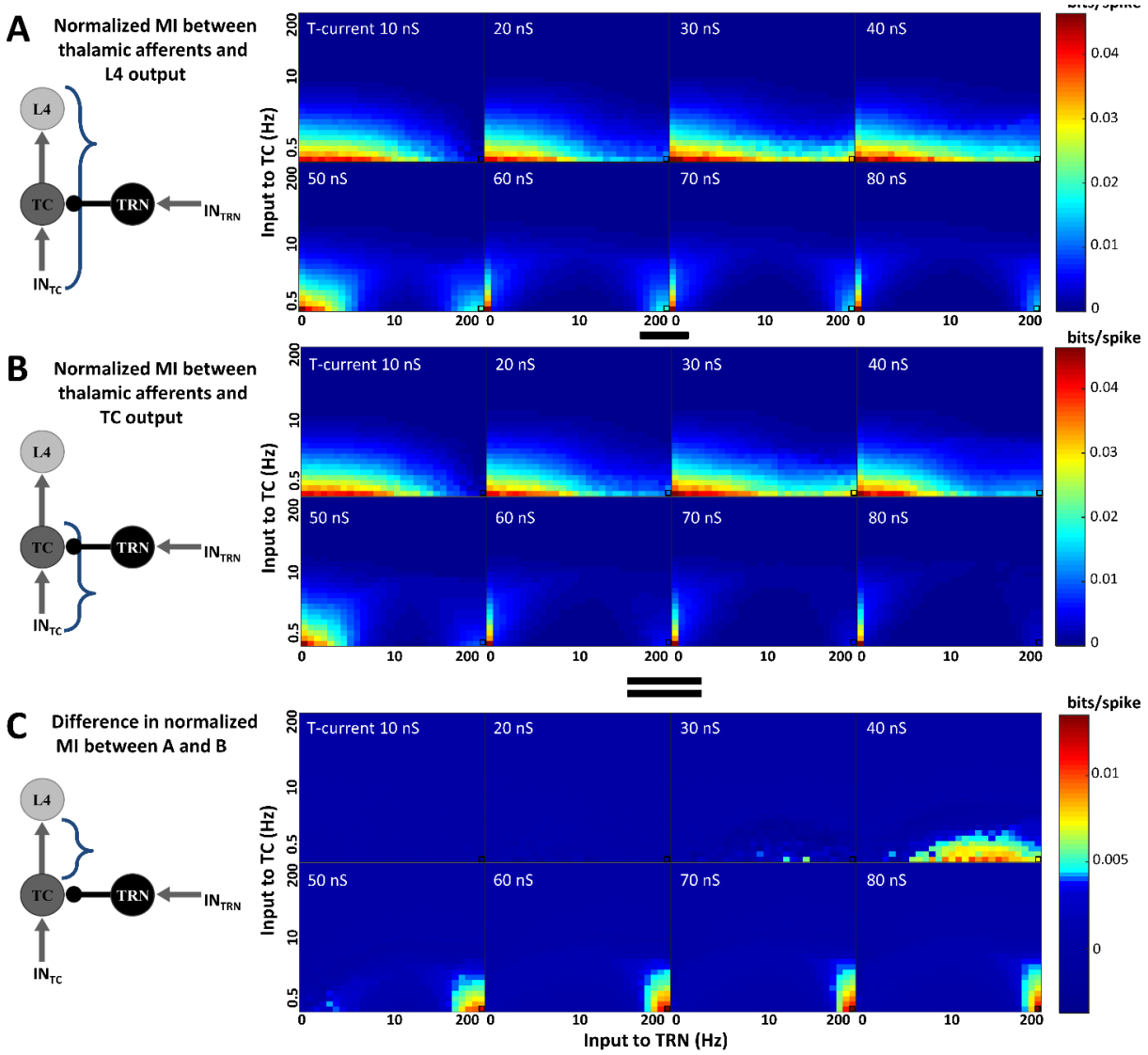

Figure 4.8. Heat map plots of normalized MI in open-loop thalamocortical network model for variable T-current conductances. For each TC T-current conductance, an MI plot is generated using a range of combinations of TC and TRN stimulation rates and then normalized by the number of output spikes for each rate combination. Normalized MI plots are averaged over 10 trials. The rate combination of 207 Hz stimulation of TRN and 0.5 Hz stimulation of TC (see **Figure 4.9**) is marked in a black box on each plot. Note that the horizontal black line between panels A and B denotes a minus sign, while the pair of horizontal black lines between panels B and C denotes an equals sign. **A)** Plots of average normalized MI (bits per output spike) between thalamic afferents and L4 output for T-current conductances of 10 nS to 80 nS. **B)** Plots of average normalized MI between thalamic afferents and TC output for T-current conductances of 10 nS to 80 nS. **C)** Plots of average normalized MI between thalamic afferents and L4 output (Fig 8A) minus average normalized MI between thalamic afferents and TC output (**Figure 4.8B**) for T-current conductances of 10 nS to 80 nS. These difference plots show a recovery of information per spike at low thalamic afferent rates and high TRN stimulation rates.

Also, for each value of T-current conductance, ranging from 10 to 80 nS, a heat map plot of normalized MI values with TRN stimulation rates on the x-axis and TC stimulation rates on the y-axis was constructed (**Figure 4.8A** for normalized MI between thalamic afferents and L4 output, and **Figure 4.8B** for normalized MI between thalamic afferents and TC output). 10 trials were performed for each data point of T-current conductance.

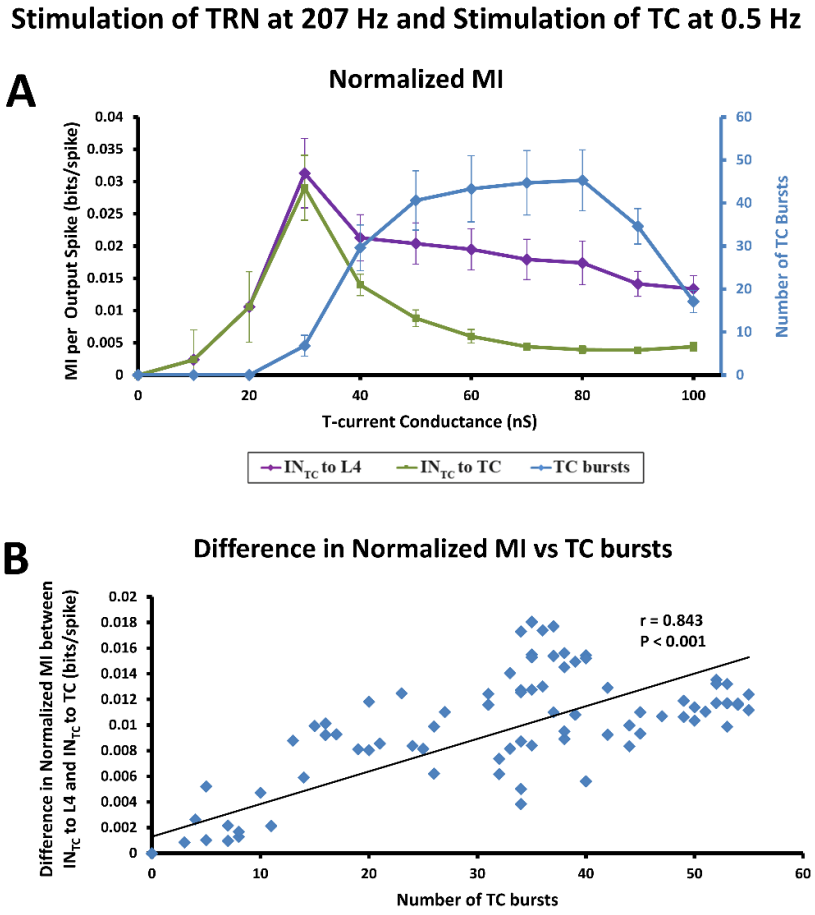

Figure 4.9. Stimulation of TRN at 207 Hz and TC at 0.5 Hz in open-loop thalamocortical network model. This rate combination is marked as a black box on the normalized MI plots of Figure 4.8. INTC refers to thalamic afferents, as shown in Figure 4.8. **A)** MI per output spike between thalamic afferents and TC output and thalamic afferents and L4 output at a range of T-current conductances from 0 nS to 100 nS. The number of bursts observed at the TC is also shown alongside on the right vertical axis. 10 trials were used for each data point and for generation of standard deviation for error bars. **B)** Difference of MI per spike of TC input to TC output and MI per spike of TC input and L4 output against the number of bursts produced by TC. The trendline is shown alongside a Pearson coefficient value of $r = 0.843$ ($p < 0.001$).

The average MI per output spike between the thalamic afferents and L4 is on average slightly greater than between the thalamic afferents and the TC cell (**Figure 4.7B**). The significance of this difference was examined by fitting each trial, consisting of MI per output spike values over all T-current conductances, to a logistic function. The average values for the coefficients a (difference between minimum and maximum value), b (steepness of the logistic curve), c (the T-current conductance midpoint of the curve), and d (the minimum value of the logistic curve) are as follows: 2.93 ± 0.17 (SD), 1.79 ± 2.86, 50.22 ± 0.29, and 1.25 ± 0.02 respectively for logistic curves fitted for the MI per output spike between the thalamic afferents and L4. The corresponding values for logistic curves fitted for the MI between thalamic afferents and the TC cell were 3.35 ± 0.16 (SD), 0.23 ± 0.01, 48.88 ± 0.25, and 0.74 ± 0.01. The values for these two sets of coefficients were found to be significantly different across all corresponding pairwise comparisons (Kruskal-Wallis and Nemenyi tests, $p < 0.001$). Note that at certain combinations of TRN and TC stimulation rates, specifically at high TRN stimulation frequencies and low TC stimulation frequencies, MI per output spike between the thalamic afferents and L4 output paradoxically becomes greater than MI per output spike between the thalamic afferents and TC output, as seen in the normalized MI difference plots of **Figure 4.8C**. For example, for TC T-current conductances above 20 nS and thalamic afferent input at 0.5 Hz and to the TRN neuron at 207 Hz, MI per spike across the whole network (from thalamic afferent to L4; purple line, **Figure 4.9A**) paradoxically exceeds that seen during the first stage of the network (from thalamic afferent to TC; green line, **Figure 4.9A**). In addition, this paradoxical behavior is seen as the degree of bursting of the TC neuron increases (blue line, **Figure 4.9A**). This relationship is quantified in **Figure 4.9B**. Here, the number of bursts in the TC neuron is compared against this paradoxical increase in MI across the whole network, and a strong positive correlation is

observed (**Figure 4.9B**, (Pearson's $r = 0.843$, $p < 0.001$)). These data suggest that the paradoxical increase in MI per spike is related to the underlying bursting behavior of the TC cell.

DISCUSSION:

Summary of findings

In the current study, the properties of a novel estimator of MI, the Adaptive partition using Interspike intervals MI Estimator (AIMIE), were examined, with its performance compared to other, more established methods of measuring MI between spike trains. The AIMIE method of computing MI was found to be insensitive to overall time compression or expansion of the spike trains, suggesting that it may be used effectively at both high and low spike rates. AIMIE was also found to be sensitive to manual manipulation of the relationship between simulated input and output spike trains. That is, manual degradation in the relationships between two spike trains also lowered the MI, computed with AIMIE, but not with other methods. Further, in paired neuronal recordings, the decrease in MI with increasing concentration of DNQX was consistent with DNQX's effect of desynchronizing the responses of the recorded neuron pairs. Finally, when AIMIE was used to compute the MI between input (to a model TC cell) and output (from a model L4 neuron) in a thalamocortical model, it was found that MI was maximum at T-current amplitudes corresponding to those seen physiologically, and that the apparent degradation in MI caused by bursting could be recovered at the thalamocortical synapse. These data suggest that a new method for estimating MI between spike trains, AIMIE, is now available to investigators and that the method has the advantage of being able to easily compensate for wide changes in the average spike rate of either the input or output spike train. In particular, this implies that AIMIE can be effectively used with neural spike data and neural networks, such as the thalamocortical network, that demonstrate high variability in spikes rate as

well as different modes of firing across different neurons. While AIMIE has not been applied in analysis of closed-loop thalamocortical network, our results suggest that AIMIE would be well-suited for this application and could possibly be used to examine the functional difference of bursting in closed-loop thalamocortical circuits and open-loop thalamocortical circuits. Thus, AIMIE may help provide further insight into how thalamocortical bursting affects relay of sensory information during awake states and during sleep states.

Methodological considerations

Since the thalamocortical network demonstrates short latencies in L4 neurons in response to afferent input provided to TC neurons (< 10 ms, corresponding to 2 synapses between input and output, (Llano et al., 2014), no effort was made to shift the output time series relative to the input time series. However, in other networks, particularly in large-scale networks containing many more synapses, where response latencies are longer, a constant time delay may need to be compensated for in the analysis. In addition, most MI estimators require a minimum amount of data for reliable results (Paninski, 2003). As observed in **Figure 4.3B-C**, under the specified simulation parameters, approximately 400-500 output spikes, corresponding to 100 seconds simulation time, are needed for AIMIE to provide consistent MI values, suggesting that for spike rates of approximately 10 spikes/second, approximately 40-50 seconds of data are required to compute MI using AIMIE. Parametric density estimators, such as maximum likelihood estimation, require relatively little data for convergence (Marek & Tichavsky, 2008), while MI estimators with uniform partition, especially DMIE, usually require larger amounts of data in comparison to MI estimators with adaptive partition (Borst & Theunissen, 1999; Walters-Williams & Li, 2009). When FBNSE and AIMIE are compared, AIMIE tends to converge faster in its estimate of MI, suggesting that its bias may be reduced more quickly than with FBNSE as

greater amounts of data are employed (**Figure 4.3D**). AIMIE shows convergence rates that are comparable to those of DMIE10 and DMIE20, while SQRSE and FBWSE show the quickest convergence rates (**Figure 4.3D**). Thus, when applying AIMIE with small enough data samples, it may be necessary to use a method for correcting sampling bias, such as the Panzeri-Treves method (Panzeri et al., 2007; Panzeri & Treves, 1996).

MI as an estimate of dependence

Multiple metrics have been used previously to compute the degree of similarity between spike trains. MI is a quantity that measures general dependence between two random variables, and as such, it is related to correlation functions which traditionally measure linear dependence (Li, 1990). STTC, for example, uses a fixed time bin around each input spike to calculate correlation between input and output spike trains (Cutts & Eglen, 2014). This approach is reminiscent of a uniform partition method, which is inadequate for analyzing MI in a TC network because of the wide range of input rates processed by TC neurons, discussed below. Related to the correlation measures are distance metrics, including the Victor-Purpura spike train distance metric (Victor & Purpura, 1996). This metric is commonly used in input and output spike train analysis, though there is evidence that this may not be an optimal metric for evaluating temporal information (Gai & Carney, 2008). Specifically, the Victor-Purpura spike train distance metric measures information carried primarily by absolute spike times, and it does not appear to effectively account for other temporal features in spike time series, such as synchronization to frequency of sensory input in a neuron's response. This insensitivity to certain temporal features could potentially complicate its use in analysis of the temporally sensitive thalamocortical network.

Furthermore, traditional correlation analysis, such as Pearson correlation, has been shown to be optimal only in the case of linear dependencies between variables, while MI can be applied to random variables that exhibit linear as well as nonlinear dependencies (Gencaga et al., 2014), which are often demonstrated in neural networks (Schöner & Kelso, 1988). Comparisons of marginal entropy, meaning the average amount of information provided by, in this case, a single spike time series, have also been used to measure the relative information content of time series in the analysis of bursting and tonic firing in the thalamus (Reinagel et al., 1999). However, only marginal entropies are compared in this earlier analysis, and there is no calculation of joint entropy between two time series, as there is in MI estimation. A consequence of not calculating joint entropy is that the relative timing between input and output spikes is lost.

Many previously employed MI estimators use uniform partitions of data to estimate density distributions (e.g., DMIE, FBNSE, FBWSE, and SQRSE). While computationally efficient, these uniform partitions can produce many empty bins (Gencaga et al., 2014). Therefore, for these estimators, different choices of bin construction for a nonadaptive partition can produce substantially different MI values. Estimators with a fixed bin width that do not depend on the number of data points, FBWSE and DMIE, show a change in MI when applied to time-scaled time series (**Figure 4.4**), unless bin widths of the partition are scaled appropriately. Furthermore, estimators with a fixed bin width that do depend on the number of data points, FBNSE and SQRSE, did not perform as well as AIMIE with certain pairs of time series of variable degrees of correlation and synchrony (**Figure 4.5** and **Figure 4.6**). These findings regarding estimators that use fixed bin widths indicate that in a network where spike timing can be changed significantly by varying input rate, an MI estimate of each time series pair would require a readjustment of the uniform partition. The use of multiple partition schemes could

cause substantial uncertainty in comparisons across samples, since every combination of TRN and TC stimulation rates would require its own parameters for the uniform partition. Therefore, for the thalamocortical network, MI estimators with uniform partitions would not be an optimal choice, leaving us to favor an adaptive partition instead. In addition, unlike estimators that use a uniform partition, AIMIE's adaptive partition also causes estimated MI to converge faster with sample size (Walters-Williams & Li, 2009), and allows investigators to work with time processes that have incomparable time-scales (Darbellay & Vajda, 1999).

T-current conductance, bursting, and MI

Systematic variation of the TC cell's T-current conductance in the thalamocortical model revealed that there is a specific range of maximal conductance values near 40 nS that produces a maximum average MI per output spike between thalamic afferent input and output at L4 (**Figure 4.7B**). This finding suggests that the TRN is able to induce the most potentiation of the ascending TC input at this peak TC T-current conductance value. Note that this value is close to the physiological TC T-current conductance of approximately 45 nS (Willis et al., 2015), and is within the range used in previous modeling studies (Deleuze et al., 2012; A. Destexhe et al., 1993; Pospischil et al., 2008; Wang, 1994). It has been shown that some hyperpolarization-activated cation currents can be modulated by factors such as intracellular pH and specific neurotransmitters, including norepinephrine and serotonin (Munsch & Pape, 1999; Pape & McCormick, 1989). If TC T-current is similarly modulated (for example see (Joksovic et al., 2010)), then the current study suggests that this modulation could alter the amount of MI between input and output as well. Thus, findings from the current study may indicate the possibility that modulation of TC T-current is one way in which the thalamocortical network can modulate sensory thalamocortical information flow.

While bursting in thalamocortical cells is observed frequently during slow-wave sleep, in which there is little relay of sensory information, it is also observed, though less frequently, during awake states in animals. Some studies suggest that thalamic cells, in particular most cells in the lateral geniculate nucleus, burst rhythmically, if only for a few seconds, during slow-wave sleep (Weyand et al., 2001). In contrast, thalamic bursting during awake states appears to show less rhythmicity, and still allows for relay of sensory information. For instance, from several experiments in visual and vibrissal systems, it is suggested that thalamic bursting in awake animals may serve to establish a more attentive and aware state, in which detection of afferent stimuli is optimized (Fanselow et al., 2001; Sherman, 2001). Previous studies have suggested that bursting in thalamic relay neurons degrades ascending sensory signals, and that tonic firing modes are more likely to produce a high-fidelity representation of ascending information en route to the cortex (Castro-Alamancos, 2002; McCarley et al., 1983; D. McCormick & Feeser, 1990). However, it has also been shown that thalamic bursts may carry more information, can enhance detection of specific temporal sequences, and are higher in synaptic efficacy than single thalamic spikes (Nicholas A. Lesica & Stanley, 2004; Nicholas A Lesica et al., 2006; Reinagel et al., 1999; H. Swadlow & A. Gusev, 2001). The current study suggests that in certain cases, particularly with high-frequency input to TRN and low frequency input to TC (**Figure 4.9A**), bursting is responsible for loss of incoming sensory information at the level of the TC (**Figure 4.9B**), as it likely obscures the ascending sensory signal. However, this information loss may be compensated in the L4 neuron due to the filter properties of the TC synapse, which produce single spikes in response to the initial spikes of a burst (Boudreau & Ferster, 2005; S. Chung et al., 2002; Gil et al., 1999; Krahe & Gabbiani, 2004; H. A. Swadlow & A. G. Gusev, 2001). Thus, information per spike that is lost at the level of the thalamus due to bursting may be partially

recovered at the level of the cortex, thus permitting perseveration of the temporal structure of the sensory stimulus while thalamic neurons are hyperpolarized by the TRN. These data suggest that thalamic bursting may coexist with high-fidelity representations in the sensory cortex: this is a hypothesis that can be tested physiologically in subsequent studies.

ACKNOWLEDGEMENTS AND CONTRIBUTIONS:

Ekaterina D. Gribkova (EDG), Baher A Ibrahim (BAEI), and Daniel A. Llano (DAL) conceived and designed research; EDG, BAEI, and DAL performed experiments; EDG, BAEI, and DAL analyzed data; EDG, BAEI, and DAL interpreted results of experiments; EDG prepared figures; EDG, BAEI, and DAL drafted manuscript; EDG, BAEI, and DAL edited and revised manuscript; EDG, BAEI, and DAL approved final version of manuscript. This work was supported by DC013073 and DC014765 to DAL. The authors thank Drs. Rama Ratnam and Jeffrey Brown (University of Illinois) for reviews of this manuscript before submission.

CHAPTER 5: CONCLUSIONS

In this thesis, I introduced and explored several different computational models inspired by neurobiological principles. The primary goals for these models are to create opportunities for enhancing AI agents and applications, and provide potential insights into biological systems, their possible evolution, and how complexity can arise from interactions of relatively simple computational elements. Motivated behavior in AI agents was explored in **Chapter 2**, with the agent-based foraging simulation, ASIMOV (Ekaterina D Gribkova et al., 2020), demonstrating how approach-avoidance decision-making in foraging can become more dynamic with the addition of relatively simple homeostatic mechanisms. While the core of ASIMOV is based upon neuronal relations used in cost-benefit choices of foraging by the predatory sea-slug *Pleurobranchaea* (Brown et al., 2018), incorporation of a small reward circuit with homeostatic plasticity resulted in the emergence of addiction processes and simple aesthetics. Complexity of behavior in both AI agents and animals is clearly enhanced with intricate memory systems, therefore, the second section of **Chapter 3** presents a model for simple episodic and sequence memory, FAM, which was integrated into ASIMOV's decision circuitry. The FAM's hippocampal-like architecture and its learning rules enabled ASIMOV's agent to learn sequences of stimuli, perform higher-order associative learning and navigate around a simple environment to find and maximize reward. A crucial element of sequence formation in the FAM is a learning rule that can determine the relative timing or order between two inputs. While hippocampal learning, such as place-cell sequence formation, clearly depends on the temporal sequence of presented stimuli, it is unclear if a single neuron is capable of learning and altering the timing of its input signals. The third section of **Chapter 3** explores this idea further, hypothesizing that a single synapse is able to learn the time delay between two signals through activity-dependent

changes in its NMDAR expression of different subunits. In particular, this synaptic model of learning, SITDL, modulates NMDARs' glutamate gate activation time to minimize mismatch between NMDAR glutamate and voltage gate conductances. With successful minimization, the resulting glutamate gate activation time effectively encodes the timing difference between synaptic glutamate and dendritic voltage signals. While this is quite useful for achieving more precise timing in systems where dendritic delays are small, SITDL may also have further importance in memory formation. I showed that with multiple synapses receiving periodic signals, SITDL mechanisms and synaptic elimination allowed for learning and reconstruction of the original glutamate signal. SITDL has significant potential for use in ANNs and machine learning, and there may be several possibilities for exploring its potential existence in physiological neurons. Overall, ASIMOV, FAM, and SITDL are relatively simple computational models that can produce notable elaboration of behavioral complexity, learning and memory in artificial agents and networks. Further, these models may also provide insights into the biological processes of episodic memory, aesthetics, and complex cognitive processes, as well as their evolution.

REFERENCES

Altier, N., & Stewart, J. (1999). The role of dopamine in the nucleus accumbens in analgesia. *Life Sciences, 65*(22), 2269-2287.

Amato, L., Davoli, M., Ferri, M., Gowing, L., & Perucci, C. A. (2004). Effectiveness of interventions on opiate withdrawal treatment: an overview of systematic reviews. *Drug and Alcohol Dependence, 73*(3), 219-226.

Ambrose, R. E., Pfeiffer, B. E., & Foster, D. J. (2016). Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron, 91*(5), 1124-1136.

Arkin, R. C. (2003). Moving up the food chain: Motivation and Emotion in behavior-based robots.

Ashida, G., & Carr, C. E. (2011). Sound localization: Jeffress and beyond. *Current Opinion in Neurobiology, 21*(5), 745-751.

Bannerman, D. M., Niewoehner, B., Lyon, L., Romberg, C., Schmitt, W. B., Taylor, A., . . . Seeburg, P. H. (2008). NMDA receptor subunit NR2A is required for rapidly acquired spatial working memory but not incremental spatial reference memory. *Journal of Neuroscience, 28*(14), 3623-3630.

Bannerman, D. M., Sprengel, R., Sanderson, D. J., McHugh, S. B., Rawlins, J. N. P., Monyer, H., & Seeburg, P. H. (2014). Hippocampal synaptic plasticity, spatial memory and anxiety. *Nature Reviews Neuroscience, 15*(3), 181-192.

Barack, D. L., & Platt, M. L. (2017). Engaging and exploring: cortical circuits for adaptive foraging decisions *Impulsivity* (pp. 163-199): Springer.

Barrera, A., Cáceres, A., Weitzenfeld, A., & Ramirez-Amaya, V. (2011). Comparative experimental studies on spatial memory and learning in rats and robots. *Journal of Intelligent & Robotic Systems, 63*(3-4), 361-397.

Barron, A. B., Maleszka, R., Helliwell, P. G., & Robinson, G. E. (2009). Effects of cocaine on honey bee dance behaviour. *Journal of Experimental Biology, 212*(2), 163-168.

Bartlett, E. L., & Wang, X. (2007). Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *Journal of Neurophysiology, 97*(2), 1005-1017.

Bartol Jr, T. M., Bromer, C., Kinney, J., Chirillo, M. A., Bourne, J. N., Harris, K. M., & Sejnowski, T. J. (2015). Nanoconnectomic upper bound on the variability of synaptic plasticity. *Elife, 4*, e10778.

Bekkers, J., & Stevens, C. (1990). *Computational implications of NMDA receptor channels.* Paper presented at the Cold Spring Harbor Symposia on Quantitative Biology.

Bellone, C., & Nicoll, R. A. (2007). Rapid bidirectional switching of synaptic NMDA receptors. *Neuron, 55*(5), 779-785.

Berlyne, D. E. (1971). *Aesthetics and psychobiology* (Vol. 336): Appleton-Century-Crofts New York.

Berridge, K. C., & Robinson, T. E. (2016). Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist, 71*(8), 670.

Bethus, I., Tse, D., & Morris, R. G. (2010). Dopamine and memory: modulation of the persistence of memory for novel hippocampal NMDA receptor-dependent paired associates. *Journal of Neuroscience, 30*(5), 1610-1618.

Blasco-Fontecilla, H., Fernández-Fernández, R., Colino, L., Fajardo, L., Perteguer-Barrio, R., & De Leon, J. (2016). The addictive model of self-harming (non-suicidal and suicidal) behavior. *Frontiers in psychiatry, 7*, 8.

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience, 2*(11), 947-957.

Boudreau, C. E., & Ferster, D. (2005). Short-term depression in thalamocortical synapses of cat primary visual cortex. *Journal of Neuroscience, 25*(31), 7179-7190.

Brown, J. W., Caetano-Anollés, D., Catanho, M., Gribkova, E., Ryckman, N., Tian, K., . . . Gillette, R. (2018). Implementing Goal-Directed Foraging Decisions of a Simpler Nervous System in Simulation. *eNeuro, 5*(1), ENEURO. 0400-0417.2018.

Burkett, J. P., & Young, L. J. (2012). The behavioral, anatomical and pharmacological parallels between social attachment, love and addiction. *Psychopharmacology, 224*(1), 1-26.

Buzsáki, G. (1989). Two-stage model of memory trace formation: a role for "noisy" brain states. *Neuroscience, 31*(3), 551-570.

Cabral, H. O., Vinck, M., Fouquet, C., Pennartz, C. M., Rondi-Reig, L., & Battaglia, F. P. (2014). Oscillatory dynamics and place field maps reflect hippocampal ensemble processing of sequence and place memory under NMDA receptor control. *Neuron, 81*(2), 402-415.

Carr, C. E. (1986). Time coding in electric fish and barn owls. *Brain, Behavior and Evolution, 28*(1-3), 122-133.

Carter, K., Lukowiak, K., Schenk, J. O., & Sorg, B. A. (2006). Repeated cocaine effects on learning, memory and extinction in the pond snail Lymnaea stagnalis. *Journal of Experimental Biology, 209*(21), 4273-4282.

Cassidy, R. M., & Tong, Q. (2017). Hunger and satiety gauge reward sensitivity. *Frontiers in Endocrinology, 8*, 104.

Castro-Alamancos, M. A. (2002). Properties of primary sensory (lemniscal) synapses in the ventrobasal thalamus and the relay of high-frequency sensory inputs. *Journal of Neurophysiology, 87*(2), 946-953.

Cellucci, C. J., Albano, A. M., & Rapp, P. (2005). Statistical validation of mutual information calculations: Comparison of alternative numerical algorithms. *Physical review E, 71*(6), 66208.

Chen, C., & Regehr, W. G. (2003). Presynaptic modulation of the retinogeniculate synapse. *The Journal of neuroscience, 23*(8), 3130-3135.

Chung, H. J., Huang, Y. H., Lau, L.-F., & Huganir, R. L. (2004). Regulation of the NMDA receptor complex and trafficking by activity-dependent phosphorylation of the NR2B subunit PDZ ligand. *Journal of Neuroscience, 24*(45), 10248-10259.

Chung, S., Li, X., & Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. *Neuron, 34*(3), 437-446.

Cooke, S., & Bliss, T. (2006). Plasticity in the human central nervous system. *Brain, 129*(7), 1659-1673.

Crabtree, J. W., & Isaac, J. T. R. (2002). New Intrathalamic Pathways Allowing Modality-Related and Cross-Modality Switching in the Dorsal Thalamus. *The Journal of Neuroscience, 22*(19), 8754-8761.

Cruikshank, S. J., Rose, H. J., & Metherate, R. (2002). Auditory thalamocortical synaptic transmission in vitro. *Journal of Neurophysiology, 87*(1), 361-384.

Cutsuridis, V., & Wennekers, T. (2009). Hippocampus, microcircuits and associative memory. *Neural Networks, 22*(8), 1120-1128.

Cutts, C. S., & Eglen, S. J. (2014). Detecting pairwise correlations in spike trains: an objective comparison of methods and application to the study of retinal waves. *The Journal of Neuroscience, 34*(43), 14288-14303.

Darbellay, G. A., & Vajda, I. (1999). Estimation of the information by an adaptive partitioning of the observation space. *IEEE Transactions on Information Theory, 45*(4), 1315-1321.

Darwin, C. (1888). *The descent of man and selection in relation to sex* (Vol. 1): Murray.

Dayan, P., & Abbott, L. F. (2001). Theoretical neuroscience: computational and mathematical modeling of neural systems.

Deleuze, C., David, F., Béhuret, S., Sadoc, G., Shin, H.-S., Uebele, V. N., . . . Bal, T. (2012). T-Type Calcium Channels Consolidate Tonic Action Potential Output of Thalamic Neurons to Neocortex. *The Journal of Neuroscience, 32*(35), 12228-12236. doi:10.1523/jneurosci.1362-12.2012

Denning, K. S., & Reinagel, P. (2005). Visual Control of Burst Priming in the Anesthetized Lateral Geniculate Nucleus. *The Journal of Neuroscience, 25*(14), 3531-3538. doi:10.1523/jneurosci.4417-04.2005

Deperrois, N., Moiseeva, V., & Gutkin, B. (2018). Minimal Circuit Model of Reward Prediction Error Computations and Effects of Nicotinic Modulations. *Frontiers in neural circuits, 12*.

Destexhe, A., Contreras, D., & Steriade, M. (1998). Mechanisms Underlying the Synchronizing Action of Corticothalamic Feedback Through Inhibition of Thalamic Relay Cells. *Journal of Neurophysiology, 79*(2), 999-1016.

Destexhe, A., McCormick, D. A., & Sejnowski, T. J. (1993). A model for 8–10 Hz spindling in interconnected thalamic relay and reticularis neurons. *Biophysical Journal, 65*(6), 2473-2477. doi:10.1016/s0006-3495(93)81297-9

Destexhe, A., McCormick, D. A., & Sejnowski, T. J. (1999). Chapter 17 Thalamic and thalamocortical mechanisms underlying 3 Hz spike-and-wave discharges. In E. R. James A. Reggia & G. Dennis (Eds.), *Progress in Brain Research* (Vol. Volume 121, pp. 289-307): Elsevier.

Dezfouli, A., Piray, P., Keramati, M. M., Ekhtiari, H., Lucas, C., & Mokri, A. (2009). A neurocomputational model for cocaine addiction. *Neural Computation, 21*(10), 2869-2893.

Dissanayake, E. (2015). "Aesthetic Primitives": Fundamental Biological Elements of a Naturalistic Aesthetics. *Aisthesis. Pratiche, linguaggi e saperi dell'estetico, 8*(1), 6-24.

Doquire, G., & Verleysen, M. (2012). *A Comparison of Multivariate Mutual Information Estimators for Feature Selection.* Paper presented at the ICPRAM (1).

Dupuis, J. P., & Groc, L. (2020). Surface trafficking of neurotransmitter receptors: From cultured neurons to intact brain preparations. *Neuropharmacology, 169*, 107642.

Einolghozati, A., Sardari, M., Beirami, A., & Fekri, F. (2011). *Capacity of discrete molecular diffusion channels.* Paper presented at the 2011 IEEE International Symposium on Information Theory Proceedings.

Elman, I., & Borsook, D. (2016). Common brain mechanisms of chronic pain and addiction. *Neuron, 89*(1), 11-36.

Entler, B. V., Cannon, J. T., & Seid, M. A. (2016). Morphine addiction in ants: a new model for self-administration and neurochemical analysis. *Journal of Experimental Biology, 219*(18), 2865-2869.

Fanselow, E. E., Sameshima, K., Baccala, L. A., & Nicolelis, M. A. (2001). Thalamic bursting in rats during different awake behavioral states. *Proceedings of the National Academy of Sciences, 98*(26), 15330-15335.

Futai, K., Okada, M., Matsuyama, K., & Takahashi, T. (2001). High-fidelity transmission acquired via a developmental decrease in NMDA receptor expression at an auditory synapse. *Journal of Neuroscience, 21*(10), 3342-3349.

Gai, Y., & Carney, L. H. (2008). Statistical analyses of temporal information in auditory brainstem responses to tones in noise: correlation index and spike-distance metric. *Journal of the Association for Research in Otolaryngology, 9*(3), 373-387.

Gencaga, D., Malakar, N. K., & Lary, D. J. (2014). Survey on the estimation of mutual information methods as a measure of dependency versus correlation analysis. *AIP Conference Proceedings, 1636*(1), 80-87. doi:doi:http://dx.doi.org/10.1063/1.4903714

Gentet, L. J., & Ulrich, D. (2003). Strong, reliable and precise synaptic connections between thalamic relay cells and neurones of the nucleus reticularis in juvenile rats. *The Journal of physiology, 546*(3), 801-811.

Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual Review of Psychology, 68*, 101-128.

Gerstner, W., Kempter, R., Van Hemmen, J. L., & Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature, 383*(6595), 76-78.

Gil, Z., Connors, B. W., & Amitai, Y. (1999). Efficacy of thalamocortical and intracortical synaptic connections: quanta, innervation, and reliability. *Neuron, 23*(2), 385-397.

Gilboa, A., Sekeres, M., Moscovitch, M., & Winocur, G. (2014). Higher-order conditioning is impaired by hippocampal lesions. *Current Biology, 24*(18), 2202-2207.

Gillette, R., Huang, R.-C., Hatcher, N., & Moroz, L. L. (2000). Cost-benefit analysis potential in feeding behavior of a predatory snail by integration of hunger, taste, and pain. *Proceedings of the National Academy of Sciences, 97*(7), 3585-3590.

Gribkova, E. D., Catanho, M., & Gillette, R. (2020). Simple Aesthetic Sense and Addiction Emerge in Neural Relations of Cost-Benefit Decision in Foraging. *Scientific Reports, 10*(1), 1-11.

Gribkova, E. D., Ibrahim, B. A. E., & Llano, D. A. (2018). A novel mutual information estimator to measure spike train correlations in a model thalamocortical network. *Journal of Neurophysiology*.

Halassa, Michael M., Chen, Z., Wimmer, Ralf D., Brunetti, Philip M., Zhao, S., Zikopoulos, B., . . . Wilson, Matthew A. (2014). State-Dependent Architecture of Thalamic Reticular Subnetworks. *Cell, 158*(4), 808-821. doi:http://dx.doi.org/10.1016/j.cell.2014.06.025

Halassa, M. M., Siegle, J. H., Ritt, J. T., Ting, J. T., Feng, G., & Moore, C. I. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. *Nature Neuroscience, 14*(9), 1118-1120.

Hassel, B., & Dingledine, R. (2012). Glutamate and glutamate receptors *Basic Neurochemistry* (pp. 342-366): Elsevier.

Hasselmo, M. E., Schnell, E., & Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience, 15*(7), 5249-5262.

Hawkins, R. D., Greene, W., & Kandel, E. R. (1998). Classical conditioning, differential conditioning, and second-order conditioning of the Aplysia gill-withdrawal reflex in a simplified mantle organ preparation. *Behavioral Neuroscience, 112*(3), 636.

Heberlein, U., Tsai, L. T.-Y., Kapfhamer, D., & Lasek, A. W. (2009). Drosophila, a genetic model system to study cocaine-related behaviors: a review with focus on LIM-only proteins. *Neuropharmacology, 56*, 97-106.

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cognitive Science, 30*(1), 3-41.

Hirata, A., & Castro-Alamancos, M. A. (2006). Relief of synaptic depression produces long-term enhancement in thalamocortical networks. *Journal of Neurophysiology, 95*(4), 2479-2491.

Hiratani, N., & Fukai, T. (2018). Redundancy in synaptic connections enables neurons to learn optimally. *Proceedings of the National Academy of Sciences, 115*(29), E6871-E6879.

Huang, Y. H., Schlüter, O. M., & Dong, Y. (2011). Cocaine-induced homeostatic regulation and dysregulation of nucleus accumbens neurons. *Behavioural Brain Research, 216*(1), 9-18.

Huguenard, J. R. (1998). Neuronal circuitry of thalamocortical epilepsy and mechanisms of antiabsence drug action. *Advances in Neurology, 79*, 991-999.

Hunt, D. L., Puente, N., Grandes, P., & Castillo, P. E. (2013). Bidirectional NMDA receptor plasticity controls CA3 output and heterosynaptic metaplasticity. *Nature Neuroscience, 16*(8), 1049.

Jahr, C. E., & Stevens, C. F. (1990). A quantitative description of NMDA receptor-channel kinetic behavior. *Journal of Neuroscience, 10*(6), 1830-1837.

Joksovic, P. M., Choe, W. J., Nelson, M. T., Orestes, P., Brimelow, B. C., & Todorovic, S. M. (2010). Mechanisms of Inhibition of T-Type Calcium Current in the Reticular Thalamic Neurons by 1-Octanol: Implication of the Protein Kinase C Pathway. *Molecular Pharmacology, 77*(1), 87-94. doi:10.1124/mol.109.059931

Jones, E. G. (2007). *The Thalamus*: Cambridge University Press.

Joshi, I., Yang, Y.-M., & Wang, L.-Y. (2007). Coincident activation of metabotropic glutamate receptors and NMDA receptors (NMDARs) downregulates perisynaptic/extrasynaptic NMDARs and enhances high-fidelity neurotransmission at the developing calyx of Held synapse. *Journal of Neuroscience, 27*(37), 9989-9999.

Jung, M. W., Lee, H., Jeong, Y., Lee, J. W., & Lee, I. (2018). Remembering rewarding futures: A simulation-selection model of the hippocampus. *Hippocampus, 28*(12), 913-930.

Jung, S.-C., Kim, J., & Hoffman, D. A. (2008). Rapid, bidirectional remodeling of synaptic NMDA receptor subunit composition by A-type K+ channel activity in hippocampal CA1 pyramidal neurons. *Neuron, 60*(4), 657-671.

Kamioka, H., Maeda, E., Jimbo, Y., Robinson, H. P., & Kawana, A. (1996). Spontaneous periodic synchronized bursting during formation of mature patterns of connections in cortical cultures. *Neuroscience Letters, 206*(2-3), 109-112.

Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., . . . Jones, T. R. (2015). Saturated reconstruction of a volume of neocortex. *Cell, 162*(3), 648-661.

Kawasaki, M., Rose, G., & Heiligenberg, W. (1988). Temporal hyperacuity in single neurons of electric fish. *Nature, 336*(6195), 173-176.

Kayser, A., Priebe, N. J., & Miller, K. D. (2001). Contrast-dependent nonlinearities arise locally in a model of contrast-invariant orientation tuning. *Journal of Neurophysiology, 85*(5), 2130-2149.

Kennedy, M. B., Beale, H. C., Carlisle, H. J., & Washburn, L. R. (2005). Integration of biochemical signalling in spines. *Nature Reviews Neuroscience, 6*(6), 423-434.

Kentros, C., Hargreaves, E., Hawkins, R. D., Kandel, E. R., Shapiro, M., & Muller, R. V. (1998). Abolition of long-term stability of new hippocampal place cell maps by NMDA receptor blockade. *Science, 280*(5372), 2121-2126.

Keramati, M., Ahmed, S. H., & Gutkin, B. S. (2017). Misdeed of the need: towards computational accounts of transition to addiction. *Current Opinion in Neurobiology, 46*, 142-153.

Kim, U., & McCormick, D. A. (1998). The functional influence of burst and tonic firing mode on synaptic interactions in the thalamus. *Journal of Neuroscience, 18*(22), 9500-9516.

Kim, U., Sanchez-Vives, M. V., & McCormick, D. A. (1997). Functional dynamics of GABAergic inhibition in the thalamus. *Science, 278*(5335), 130-134.

Kimura, A. (2014). Diverse subthreshold cross-modal sensory interactions in the thalamic reticular nucleus: implications for new pathways of cross-modal attentional gating function. *European Journal of Neuroscience, 39*(9), 1405-1418.

Kjaer, T. W., Hertz, J. A., & Richmond, B. J. (1994). Decoding cortical neuronal signals: network models, information estimation and spatial tuning. *Journal of Computational Neuroscience, 1*(1-2), 109-139.

Kohonen, T. (1972). Correlation matrix memories. *IEEE Transactions on Computers, 100*(4), 353-359.

Krahe, R., & Gabbiani, F. (2004). Burst firing in sensory systems. *Nature Reviews Neuroscience, 5*(1), 13-23.

Kunec, S., Hasselmo, M. E., & Kopell, N. (2005). Encoding and retrieval in the CA3 region of the hippocampus: a model of theta-phase separation. *Journal of Neurophysiology, 94*(1), 70-82.

Kusayama, T., & Watanabe, S. (2000). Reinforcing effects of methamphetamine in planarians. *NeuroReport: For Rapid Communication of Neuroscience Research, 11*(11), 2511-2513.

Laurent, A., Goaillard, J.-M., Cases, O., Lebrand, C., Gaspar, P., & Ropert, N. (2002). Activity-dependent presynaptic effect of serotonin 1B receptors on the somatosensory thalamocortical transmission in neonatal mice. *The Journal of neuroscience, 22*(3), 886-900.

Lawrence, M., Trappenberg, T., & Fine, A. (2006). Rapid learning and robust recall of long sequences in modular associator networks. *Neurocomputing, 69*(7-9), 634-641.

Lee, H.-G., Kim, Y.-C., Dunning, J. S., & Han, K.-A. (2008). Recurring ethanol exposure induces disinhibited courtship in Drosophila. *PloS One, 3*(1), e1391.

Leknes, S., & Tracey, I. (2008). A common neurobiology for pain and pleasure. *Nature Reviews Neuroscience, 9*(4), 314.

Leri, F., Bruneau, J., & Stewart, J. (2003). Understanding polydrug use: Review of heroin and cocaine co-use. *Addiction, 98*(1), 7-22.

Lesica, N. A., & Stanley, G. B. (2004). Encoding of Natural Scene Movies by Tonic and Burst Spikes in the Lateral Geniculate Nucleus. *The Journal of Neuroscience, 24*(47), 10731-10740. doi:10.1523/jneurosci.3059-04.2004

Lesica, N. A., Weng, C., Jin, J., Yeh, C.-I., Alonso, J.-M., & Stanley, G. B. (2006). Dynamic encoding of natural luminance sequences by LGN bursts. *PLoS Biology, 4*(7), e209.

Lester, R. A., Clements, J. D., Westbrook, G. L., & Jahr, C. E. (1990). Channel kinetics determine the time course of NMDA receptor-mediated synaptic currents. *Nature, 346*(6284), 565-567.

Li, W. (1990). Mutual information functions versus correlation functions. *Journal of statistical physics, 60*(5-6), 823-837.

Lippe, W. (1994). Rhythmic spontaneous activity in the developing avian auditory system. *Journal of Neuroscience, 14*(3), 1486-1495.

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*.

Lisman, J. E. (1997). Bursts as a unit of neural information: making unreliable synapses reliable. *Trends in Neurosciences, 20*(1), 38-43.

Llano, D. A., Slater, B. J., Lesicko, A. M., & Stebbings, K. A. (2014). An auditory colliculothalamocortical brain slice preparation in mouse. *Journal of Neurophysiology, 111*(1), 197-207.

Llinás, R. R., & Steriade, M. (2006). Bursting of thalamic neurons and states of vigilance. *Journal of Neurophysiology, 95*(6), 3297-3308.

Lüscher, C., & Malenka, R. C. (2012). NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). *Cold Spring Harbor Perspectives in Biology*, a005710.

Lutter, M., & Nestler, E. J. (2009). Homeostatic and hedonic signals interact in the regulation of food intake. *The Journal of nutrition, 139*(3), 629-632.

MacLean, J. N., Watson, B. O., Aaron, G. B., & Yuste, R. (2005). Internal dynamics determine the cortical response to thalamic stimulation. *Neuron, 48*(5), 811-823.

Madl, T., Chen, K., Montaldi, D., & Trappl, R. (2015). Computational cognitive models of spatial memory in navigation space: A review. *Neural Networks, 65*, 18-43.

Marek, T., & Tichavsky, P. (2008). *On the estimation of mutual information.* Paper presented at the Proceedings of ROBUST.

McAlonan, K., Cavanaugh, J., & Wurtz, R. H. (2008). Guarding the gateway to cortex with attention in visual thalamus. *Nature, 456*(7220), 391-394. doi:http://www.nature.com/nature/journal/v456/n7220/suppinfo/nature07382_S1.html

McCarley, R. W., Benoit, O., & Barrionuevo, G. (1983). Lateral geniculate nucleus unitary discharge in sleep and waking: state-and rate-specific aspects. *Journal of Neurophysiology, 50*(4), 798-818.

McCormick, D., & Feeser, H. (1990). Functional implications of burst firing and single spike activity in lateral geniculate relay neurons. *Neuroscience, 39*(1), 103-113.

McCormick, D. A., & Contreras, D. (2001). On the cellular and network bases of epileptic seizures. *Annual Review of Physiology, 63*(1), 815-846.

McSweeney, F. K., & Murphy, E. S. (2009). Sensitization and habituation regulate reinforcer effectiveness. *Neurobiology of Learning and Memory, 92*(2), 189-198.

Mechner, F. (2018). A Behavioral and Biological Analysis of Aesthetics: Implications for Research and Applications. *The Psychological Record, 68*(3), 287-321. doi:10.1007/s40732-017-0228-1

Miller, J.-e. K., Ayzenshtat, I., Carrillo-Reid, L., & Yuste, R. (2014). Visual stimuli recruit intrinsically generated cortical ensembles. *Proceedings of the National Academy of Sciences, 111*(38), E4053-E4061.

Mosteller, F., & Tukey, J. W. (1977). Data analysis and regression: a second course in statistics. *Addison-Wesley Series in Behavioral Science: Quantitative Methods*.

Mukherjee, P., & Kaplan, E. (1995). Dynamics of neurons in the cat lateral geniculate nucleus: in vivo electrophysiology and computational modeling. *Journal of Neurophysiology, 74*(3), 1222-1243.

Munsch, T., & Pape, H. C. (1999). Modulation of the hyperpolarization-activated cation current of rat thalamic relay neurones by intracellular pH. *The Journal of Physiology, 519*(2), 493-504.

Nathaniel, T. I., Panksepp, J., & Huber, R. (2009). Drug-seeking behavior in an invertebrate system: evidence of morphine-induced reward, extinction and reinstatement in crayfish. *Behavioural Brain Research, 197*(2), 331-338.

Navratilova, E., & Porreca, F. (2014). Reward and motivation in pain and pain relief. *Nature Neuroscience, 17*(10), 1304.

Nestler, E. J. (2005). Is there a common molecular pathway for addiction? *Nature Neuroscience, 8*(11), 1445.

Noboa, V., & Gillette, R. (2013). Selective prey avoidance learning in the predatory sea slug *Pleurobranchaea californica*. *Journal of Experimental Biology, 216*(17), 3231-3236.

Oram, M., Wiener, M., Lestienne, R., & Richmond, B. (1999). Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *Journal of Neurophysiology, 81*(6), 3021-3033.

Oram, M. W., Hatsopoulos, N. G., Richmond, B. J., & Donoghue, J. P. (2001). Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *Journal of Neurophysiology, 86*(4), 1700-1716.

Oswald, A.-M. M., Doiron, B., & Maler, L. (2007). Interval coding. I. Burst interspike intervals as indicators of stimulus intensity. *Journal of Neurophysiology, 97*(4), 2731-2743.

Palladini, G., Ruggeri, S., Stocchi, F., De Pandis, M., Venturini, G., & Margotta, V. (1996). A pharmacological study of cocaine activity in planaria. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology, 115*(1), 41-45.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Computation, 15*(6), 1191-1253. doi:10.1162/089976603321780272

Panzeri, S., Senatore, R., Montemurro, M. A., & Petersen, R. S. (2007). Correcting for the sampling bias problem in spike train information measures. *Journal of Neurophysiology, 98*(3), 1064-1072.

Panzeri, S., & Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures. *Network: Computation in Neural Systems, 7*(1), 87-107.

Paoletti, P., Bellone, C., & Zhou, Q. (2013). NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nature Reviews Neuroscience, 14*(6), 383-400.

Pape, H.-C., & McCormick, D. A. (1989). Noradrenaline and serotonin selectively modulate thalamic burst firing by enhancing a hyperpolarization-activated cation current. *Nature, 340*(6236), 715-718.

Person, A. L., & Perkel, D. J. (2005). Unitary IPSPs Drive Precise Thalamic Spiking in a Circuit Required for Learning. *Neuron, 46*(1), 129-140. doi:10.1016/j.neuron.2004.12.057

Pinault, D., & Deschênes, M. (1998). Anatomical evidence for a mechanism of lateral inhibition in the rat thalamus. *European Journal of Neuroscience, 10*(11), 3462-3469.

Ponomarenko, A., & Korotkova, T. (2018). Hunger is a gatekeeper of pain in the brain. *Nature, 556*(7702), 445.

Pospischil, M., Toledo-Rodriguez, M., Monier, C., Piwkowska, Z., Bal, T., Frégnac, Y., . . . Destexhe, A. (2008). Minimal Hodgkin–Huxley type models for different classes of cortical and thalamic neurons. *Biological Cybernetics, 99*(4-5), 427-441.

Prum, R. O. (2017). *The Evolution of Beauty: How Darwin's Forgotten Theory of Mate Choice Shapes the Animal World-and Us*: Anchor.

Ramcharan, E., Gnadt, J., & Sherman, S. (2000). Burst and tonic firing in thalamic cells of unanesthetized, behaving monkeys. *Visual Neuroscience, 17*(1), 55.

Redish, A. D. (2004). Addiction as a computational process gone awry. *Science, 306*(5703), 1944-1947.

Redish, A. D., Jensen, S., & Johnson, A. (2008). Addiction as vulnerabilities in the decision process. *Behavioral and Brain Sciences, 31*(4), 461-487.

Reinagel, P., Godwin, D., Sherman, S. M., & Koch, C. (1999). Encoding of Visual Information by LGN Bursts. *Journal of Neurophysiology, 81*(5), 2558-2569.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory, 2*, 64-99.

Rolls, E. T. (2011). The Origins of Aesthetics: A Neurobiological Basis for Affective Feelings and Aesthetics. In E. Schellekens & P. Goldie (Eds.), *The Aesthetic Mind: Philosophy and Psychology* (pp. 116–165). Oxford, USA: Oxford University Press.

Rose, H. J., & Metherate, R. (2005). Auditory thalamocortical transmission is reliable and temporally precise. *Journal of Neurophysiology, 94*(3), 2019-2030.

Rosenberg, K. P., & Feder, L. C. (2014). An introduction to behavioral addictions *Behavioral Addictions* (pp. 1-17): Elsevier.

Sadowski, J. H., Jones, M. W., & Mellor, J. R. (2011). Ripples make waves: binding structured activity and plasticity in hippocampal networks. *Neural Plasticity, 2011*.

Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A., & Silver, R. A. (2005). Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *Journal of Neuroscience, 25*(36), 8173-8187.

Sasakawa, T., Hu, J., & Hirasawa, K. (2008). A brainlike learning system with supervised, unsupervised, and reinforcement learning. *Electrical Engineering in Japan, 162*(1), 32-39.

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development, 2*(3), 230-247.

Schöner, G., & Kelso, J. A. S. (1988). Dynamic Pattern Generation in Behavioral and Neural Systems. *Science, 239*(4847), 1513-1520.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*: Cambridge university press.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*(3), 379-423. doi:10.1002/j.1538-7305.1948.tb01338.x

Sherman, S. M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *Trends in Neurosciences, 24*(2), 122-126.

Shipley, A. T., Imeh-Nathaniel, A., Orfanakos, V. B., Wormack, L. N., Huber, R., & Nathaniel, T. I. (2017). The sensitivity of the crayfish reward system to mammalian drugs of abuse. *Frontiers in Physiology, 8*, 1007.

Shouval, H. Z., Bear, M. F., & Cooper, L. N. (2002). A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences, 99*(16), 10831-10836.

Silverman, B. W. (1986) *Density estimation for statistics and data analysis* (Vol. 26): CRC press.

Sincich, L. C., Adams, D. L., Economides, J. R., & Horton, J. C. (2007). Transmission of spike trains at the retinogeniculate synapse. *The Journal of neuroscience, 27*(10), 2683-2692.

Singh, P., Doshi, S., Spaethling, J. M., Hockenberry, A. J., Patel, T. P., Geddes-Klein, D. M., . . . Meaney, D. F. (2012). N-methyl-D-aspartate receptor mechanosensitivity is governed by C terminus of NR2B subunit. *Journal of Biological Chemistry, 287*(6), 4348-4359.

Singh, P., Hockenberry, A. J., Tiruvadi, V. R., & Meaney, D. F. (2011). Computational investigation of the changing patterns of subtype specific NMDA receptor activation during physiological glutamatergic neurotransmission. *PLoS Computational Biology, 7*(6).

Smith, D. M., & Mizumori, S. J. (2006). Hippocampal place cells, context, and episodic memory. *Hippocampus, 16*(9), 716-729.

Smith, G. D., Cox, C. L., Sherman, S. M., & Rinzel, J. (2000). Fourier Analysis of Sinusoidally Driven Thalamocortical Relay Neurons and a Minimal Integrate-and-Fire-or-Burst Model. *Journal of Neurophysiology, 83*(1), 588-610.

Solomon, R. L., & Corbit, J. D. (1974). An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review, 81*(2), 119.

Spiro, J. E., Brose, N., Heinemann, S. F., & Heiligenberg, W. (1994). Immunolocalization of NMDA receptors in the central nervous system of weakly electric fish: functional implications for the modulation of a neuronal oscillator. *Journal of Neuroscience, 14*(10), 6289-6299.

Stebbings, K. A., Choi, H. W., Ravindra, A., Caspary, D. M., Turner, J. G., & Llano, D. A. (2016). Ageing-related changes in GABAergic inhibition in mouse auditory cortex, measured using in vitro flavoprotein autofluorescence imaging. *J Physiol, 594*(1), 207-221. doi:10.1113/JP271221

Steels, L. (1993). The artificial life roots of artificial intelligence. *Artificial Life, 1*(1_2), 75-110.

Steriade, M., & Llinás, R. R. (1988). The functional states of the thalamus and the associated neuronal interplay. *Physiological Reviews, 68*(3), 649-742.

Steriade, M., McCormick, D., & Sejnowski, T. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science, 262*(5134), 679-685. doi:10.1126/science.8235588

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, W. (1998). Entropy and Information in Neural Spike Trains. *Physical Review Letters, 80*(1), 197-200.

Strösslin, T., Sheynikhovich, D., Chavarriaga, R., & Gerstner, W. (2005). Robust self-localisation and navigation based on hippocampal place cells. *Neural Networks, 18*(9), 1125-1140.

Subagdja, B., & Tan, A.-H. (2015). Neural modeling of sequential inferences and learning over episodic memory. *Neurocomputing, 161*, 229-242.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*: MIT press.

Swadlow, H., & Gusev, A. (2001). The impact of 'bursting' thalamic impulses at a neocortical synapse. *Nature Neuroscience, 4*(4), 402-408.

Swadlow, H. A., & Gusev, A. G. (2001). The impact of'bursting'thalamic impulses at a neocortical synapse. *Nature Neuroscience, 4*(4), 402-408.

Takahashi, Y., Schoenbaum, G., & Niv, Y. (2008). Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. *Frontiers in Neuroscience, 2*, 14.

Tang, Y. Z., & Carr, C. E. (2007). Development of N-methyl-D-aspartate receptor subunits in avian auditory brainstem. *Journal of Comparative Neurology, 502*(3), 400-413.

Tolman, E. C. (1939). Prediction of vicarious trial and error by means of the schematic sowbug. *Psychological Review, 46*(4), 318.

Toni, N., Buchs, P.-A., Nikonenko, I., Bron, C., & Muller, D. (1999). LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. *Nature, 402*(6760), 421-425.

Turrigiano, G. G., & Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience, 5*(2), 97.

Van Strien, N., Cappaert, N., & Witter, M. (2009). The anatomy of memory: an interactive overview of the parahippocampal–hippocampal network. *Nature Reviews Neuroscience, 10*(4), 272.

Ventriglia, F., & Di Maio, V. (2000). A Brownian simulation model of glutamate synaptic diffusion in the femtosecond time scale. *Biological Cybernetics, 83*(2), 93-109.

Victor, J. D., & Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: a metric-space analysis. *Journal of Neurophysiology, 76*(2), 1310-1326.

Volkow, N. D., Michaelides, M., & Baler, R. (2019). The Neuroscience of Drug Reward and Addiction. *Physiological Reviews, 99*(4), 2115-2140.

Walters-Williams, J., & Li, Y. (2009). Estimation of Mutual Information: A Survey. In P. Wen, Y. Li, L. Polkowski, Y. Yao, S. Tsumoto, & G. Wang (Eds.), *Rough Sets and Knowledge Technology* (pp. 389-396). Berlin, Heidelberg: Springer Berlin Heidelberg.

Walther, B., Morgenstern, M., & Hanewinkel, R. (2012). Co-occurrence of addictive behaviours: personality factors related to substance use, gambling and computer gaming. *European Addiction Research, 18*(4), 167-174.

Wang, X.-J. (1994). Multiple dynamical modes of thalamic relay neurons: rhythmic bursting and intermittent phase-locking. *Neuroscience, 59*(1), 21-31.

Watson, B. O., MacLean, J. N., & Yuste, R. (2008). UP states protect ongoing cortical activity from thalamic inputs. *PloS One, 3*(12), e3971.

Weyand, T. G., Boudreaux, M., & Guido, W. (2001). Burst and tonic response modes in thalamic neurons during sleep and wakefulness. *Journal of Neurophysiology, 85*(3), 1107-1118.

Whitmire, C. J., Millard, D. C., & Stanley, G. B. (2017). Thalamic state control of cortical paired-pulse dynamics. *Journal of Neurophysiology, 117*(1), 163-177.

Whitmire, C. J., Waiblinger, C., Schwarz, C., & Stanley, G. B. (2016). Information coding through adaptive gating of synchronized thalamic bursting. *Cell reports, 14*(4), 795-807.

Wilensky, U. (1999). NetLogo: Center for connected learning and computer-based modeling. *Northwestern University, Evanston, IL, 4952*.

Willis, A. M., Slater, B. J., Gribkova, E. D., & Llano, D. A. (2015). Open-loop organization of thalamic reticular nucleus and dorsal thalamus: a computational model. *Journal of Neurophysiology, 114*(4), 2353-2367.

Wilson, I. A., Gallagher, M., Eichenbaum, H., & Tanila, H. (2006). Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends in Neurosciences, 29*(12), 662-670.

Worms, J., & Touati, S. (2016). *Parametric and Non-Parametric Statistics for Program Performance Analysis and Comparison*. Retrieved from https://hal.inria.fr/hal-01286112

Xenakis, I., & Arnellos, A. (2014). Aesthetic perception and its minimal content: a naturalistic perspective. *Frontiers in Psychology, 5*, 1038.

Young, P. T. (1946). Studies of food preference, appetite and dietary habit. VI. Habit, palatability and diet as factors regulating the selection of food by the rat. *Journal of Comparative Psychology, 39*(3), 139.

Yu, Q., Tang, H., Hu, J., & Tan, K. C. (2017). *Neuromorphic Cognitive Systems*: Springer.

Zakon, H., Oestreich, J., Tallarovic, S., & Triefenbach, F. (2002). EOD modulations of brown ghost electric fish: JARs, chirps, rises, and dips. *Journal of Physiology-Paris, 96*(5-6), 451-458.

Zhang, W., & Zhou, B. (2017). Learning to update auto-associative memory in recurrent neural networks for improving sequence memorization. *arXiv preprint arXiv:1709.06493*.

Zheng, D., de Vaca, S. C., & Carr, K. D. (2012). Food restriction increases acquisition, persistence and drug prime-induced expression of a cocaine-conditioned place preference in rats. *Pharmacology Biochemistry and Behavior, 100*(3), 538-544.

Zikopoulos, B., & Barbas, H. (2012). Pathways for Emotions and Attention Converge on the Thalamic Reticular Nucleus in Primates. *The Journal of Neuroscience, 32*(15), 5338-5350. doi:10.1523/jneurosci.4793-11.2012